EXHIBIT 17

Redacted Version of Document to be Sealed

UNITED STATES DISTRICT COURT
EASTERN DISTRICT OF MICHIGAN
SOUTHERN DIVISION

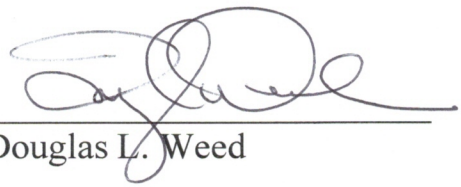
ELNORA CARTHAN, et al., Plaintiffs,	
v.	Case No. 5:16-cv-10444-JEK-MKM
RICK SNYDER, et al., Defendants.	Hon. Judith E. Levy Magistrate Judge Mona K. Majzoub

---

DECLARATION OF DOUGLAS L. WEED

I, Douglas L. Weed, M.D., M.P.H., Ph.D., submit the attached report in support of the VNA Defendants' Opposition to the Plaintiffs' Motion for Class Certification. I declare under penalty of perjury that the statements made in my report are true and accurate to the best of my information and knowledge.

1.5.2021
Date



Douglas L. Weed

Expert Report of Douglas L. Weed, M.D., M.P.H., Ph.D.
Regarding General Causation and Medical Monitoring
CLASS CERTIFICATION

Elnora Carthan et al., Plaintiffs, v. Rick Snyder et al., Defendants
Case No. 5:16-cv-10444-JEK-MKM

TABLE OF CONTENTS

PART ONE: INTRODUCTION

SECTION 1.1 QUALIFICATIONS

My name is Douglas L. Weed. I am the founder and managing member of DLW Consulting Services LLC, a scientific consulting firm specializing in epidemiology with a focus on disease causation. I am a physician-epidemiologist with over 37 years of experience in epidemiology, epidemiological methods, and the methods of causal inference.

EDUCATION

I received a Bachelor of Science (B.Sc.) degree in Engineering, *summa cum laude*, in 1974 from The Ohio State University. In 1977, I received the Doctor of Medicine (M.D.) degree from The Ohio State University. In 1980, I received the Master of Public Health (M.P.H.) degree from the University of North Carolina and the Doctor of Philosophy (Ph.D.) degree in Epidemiology from the University of North Carolina in 1982.

WORK EXPERIENCE

In 1982, after completing the Ph.D. in occupational and environmental epidemiology, I was hired by the National Cancer Institute (NCI) as a Senior Staff Fellow, a research position. I was progressively promoted to the position, Chief, Office of Preventive Oncology, a post which I held until 2007. During the years 1990 through 2007, I directed the Cancer Prevention Fellowship Program, a postdoctoral training program at the National Cancer Institute in the Division of Cancer Prevention. Note that one of the key concerns of cancer prevention is the identification of causes of cancer. I also created and directed the Summer Curriculum in Cancer Prevention and Control, a multidisciplinary educational program. Both programs provide state-of-the-art postdoctoral training in cancer etiology, prevention, and control. Since leaving the NCI in 2007, I have been engaged in the professional consulting practice of epidemiology, the scientific study of the causes and distributions of human disease and health conditions and the application of that knowledge to improve the public's health. I am now retired from the NCI.

PROFESSIONAL EXPERIENCE

Currently I am an adjunct full professor at the University of Utah's School of Medicine in the Department of Family and Preventive Medicine, Division of Public Health. Previously, I held academic positions at the Johns Hopkins University, Georgetown University, the Uniformed Services University of the Health Sciences and the University of New Mexico as well as visiting professorships at the National School of Public Health in Madrid, Spain and at McGill University in Montreal. In addition, I was a Visiting Scholar at the Federal Judicial Center in Washington D.C. and co-chaired a National Academy of Sciences (NAS) committee examining the *Daubert* decision on its 10th anniversary.

I am the author of over 140 published scientific papers. I have edited and contributed to two books, one on cancer prevention and control published in 1996 and another on ethics and epidemiology published in 2009.

I am a Fellow of the American College of Epidemiology (ACE) and formerly, Chair of the Ethics and Standards of Practice Committee of the College. As a member of that committee, I co-wrote the

American College of Epidemiology's Ethics Guidelines (2000). I have held leadership positions in the field of epidemiology, as a member of the Executive Committee of the Society for Epidemiologic Research and as a member of the Board of Directors of the American College of Epidemiology.

RESEARCH INTERESTS AND EXPERIENCE

My primary research focus is the science and practice of disease causation. Over the course of the past 37 years, I have published peer reviewed papers on the methods used to assess causation, the practice of causal inference, theories of causation, the logic of causal inference, and the philosophies of science applicable to this central problem of medicine, public health, and the law. For examples of papers on causality in the peer reviewed literature, see the following: Weed (1986), Koopman and Weed (1990), Weed and Gorelic (1996), Weed and Hursting (1998), Weed (2000a), Weed (2002), Weed (2005), Parascandola and Weed (2006), Weed (2016) and Weed (2018). I have also published peer-reviewed papers on the methods and practice of meta-analysis as applied to epidemiological studies. See, for example: Weed (2000b), Weed (2010), Alexander et al. (2011), and Althuis et al. (2014). I have also published peer-reviewed papers on the methods and practice of systematic reviews. See, for example, Breslow et al. (1998), Weed et al. (2011), Alexander et al. (2012), Weed (2013), Alexander et al. (2013), Alexander and Weed (2016) and Weed (2018). Finally, I have published analyses of mortality in occupational populations; see for example, Weed et al. (1987) and Weed (2010).

I have lectured on disease causation and epidemiological methods at the National Cancer Institute, the Institute of Medicine, the United States Environmental Protection Agency, the National Academies of Science, Harvard University, Yale University, University of California (Berkeley), Imperial College (London), the University of Michigan, the University of North Carolina, Sloan Kettering Cancer Center, MD Anderson Cancer Center, the University of New Mexico, the University of Utah, and at The Ohio State University as well as at academic and research institutions around the world (e.g. China, Japan, Germany, Ireland, Norway, and Turkey) and at many scientific conferences. Most recently, for example, I presented at the University of Alabama at Birmingham in July, 2016, at the International Society of Environmental Epidemiology meeting in Rome, Italy, September, 2016, at the American Association for Cancer Research meeting in Washington DC, April, 2017, at the International Epidemiological Association meeting in Japan, September, 2017, at the Society of Toxicology meeting in San Antonio, Texas, March, 2018, at Queens University, Belfast, Northern Ireland, September, 2018, at the National Cancer Institute in Rockville, Maryland, November, 2018, at George Washington University in Washington D.C., March, 2019, and at the University of Indiana, July, 2019.

I have trained and taught hundreds of physicians, nurses, public health scientists, and biomedical scientists in the principles and practice of disease causation. In addition, I have designed epidemiological studies, meta-analyses, and systematic reviews, analyzed the results, and published my findings in peer-reviewed journals.

I have also written extensively and lectured on bioethics, with special interest in the application of bioethical principles and methods to biomedical research, epidemiology, preventive medicine, and public health (Weed and McKeown, 2001). Topics of special interest have been scientific misconduct (Weed, 1998), the ethics of cancer screening (Weed, 1999), and conflicts of interest (Weed, 2009).

Finally, I have extensive experience in the peer review of manuscripts submitted for publication. I have been (and continue to be) asked to review manuscripts for at least 30 different scientific journals. Examples include: the Journal of the American Medical Association (JAMA), Cancer, Critical Reviews in

Toxicology, the American Journal of Public Health (AJPH), the American Journal of Epidemiology (AJE), the American Journal of Preventive Medicine, Nutrition and Cancer, Global Epidemiology, and the Journal of the National Cancer Institute (JNCI). In addition, I serve on the editorial board of the JNCI where I manage this same peer review process. I have been a Reviews Editor for JNCI for the past 22 years.

I am being compensated at the rate of $600 per hour. My curriculum vitae can be found in Appendix E. My opinions are expressed to a reasonable degree of medical and scientific certainty.

Section 1.2 Purpose of the Declaration

I have been asked to provide an opinion to a reasonable degree of scientific certainty on the following general questions:

Is class certification scientifically justified for the following issues:

1. Exposure to lead and neurodevelopmental outcomes
2. Medical Monitoring

I have also been asked to provide an opinion to a reasonable degree of scientific certainty on the following more specific questions:

Does exposure to lead in drinking water cause cognitive deficits in children?

If so, what are the scientific characteristics of the relationship as they apply to the child residents of Flint, Michigan in the context of the Flint Water Switch of 2014-15?

Does exposure to lead in drinking water cause behavioral disorders (to include aggression, hyperactivity, inattention, and impulsivity) in children?

If so, what are the scientific characteristics of the relationship as they apply to the child residents of Flint, Michigan in the context of the Flint Water Switch of 2014-15?

Has every member of the group defined by the plaintiffs as the “Minors Subclass” been harmed? By “harmed” I mean experienced a deficit in IQ and/or a diagnosis (or experienced symptoms) of attention deficit hyperactivity disorder (ADHD) caused by lead exposure from the Flint Water Switch.

Is it scientifically and ethically separately appropriate on a class basis to recommend and implement a medical monitoring program based on the concept of early detection—as proposed by the plaintiffs’ experts, Drs. Ducatman and Keating—for the group of children of Flint, Michigan defined by the plaintiffs as the “Minors Subclass?”

Are the opinions of the plaintiffs’ experts, Drs. Lanphear, Hu, Ducatman, and Keating scientifically valid and reliable?

If additional scientific issues emerge other than those described above[1], I reserve the right to supplement my report to address them.

Furthermore, if additional information (e.g., in the form of studies, reports, legal documents, etc.) becomes available after my report is submitted, I reserve the right to supplement this report in response to that information.

Section 1.3 A Caveat Concerning the Purpose of this Report

I do not address here the question of whether populations (or, for that matter, individuals) should avoid exposure to lead in drinking water whether by personal choice or governmental regulations. There is a critically important distinction to be made between scientific claims of causation and medical practice (including public health and preventive) recommendations and actions. As Sir Austin Bradford Hill, perhaps the best-known scientist (ever) who has proposed a method for determining general causation, has written (Hill, 1965, p. 300):

> "Finally, in passing from association to causation, I believe in 'real life' we shall have to consider what flows from that decision. **On scientific grounds we should do no such thing.** The evidence is there to be judged on its merits and the judgment (in that sense) should be utterly independent of what hangs upon it—or who hangs because of it." (emphasis added)

Simply put, causation questions can be (and will be here) kept separate from medical, public health, and regulatory decision-making.

I will not discuss this issue in detail. What is important to keep in mind is that there is a sharp distinction between the scientific question such as "does exposure to lead in drinking water cause cognitive disorders?" and the regulatory (or medical practice question): "should exposure to lead in drinking water (at some specific amount and duration) be avoided?"

The remainder of PART ONE of this report is an executive summary of my opinions. The focus of PART TWO of this report examines the role of life course epidemiology, i.e. factors other than lead that cause (or increase the risk of) neurodevelopmental outcomes in children. PART THREE deals with sources of lead exposure including but not limited to lead in drinking water as well as cognitive and behavioral disorders in children (referred to broadly as "neurodevelopmental disorders"). PART FOUR is a description and interpretation of published studies on blood lead levels in Flint, Michigan children before, during, and after the switch from the Detroit Water Authority to the Flint River water. PART FIVE of this report is focused solely on the scientific question of general causation. There I provide a brief description of the methods currently used by the scientific community to assess general causation as well as the application of those methods to the central issues of the hypothetical relationships

[1] Please note that several additional general causation issues are addressed in this report, beyond those described above. These issues include exposure to low levels of lead and kidney disease (and kidney dysfunction) in children, exposure to low levels of lead and increased blood pressure (hypertension) in children. Also discussed are the issues of exposure to low levels of lead as children and the occurrence of cardiovascular disease as adults as well as exposure to low levels of lead as children and the occurrence of essential tremor in adulthood. These issues are discussed in Appendices G, H, I, and J respectively. I also address the issue of lead exposure and spontaneous abortion.

between lead exposure and cognitive and behavioral outcomes as they are applied to the Flint Water Switch and the population of Flint, Michigan. PART SIX contains my assessments of the methods (if any) and opinions of the class certification plaintiffs' experts, Drs. Hu, Lanphear, Ducatman, and Keating. PART SEVEN contains my opinions. PART EIGHT contains reference lists. PART NINE contains appendices.

Section 1.4 Executive Summary I: Failure of the Class Certification Claims and Opinions

The purpose of this section is to summarize my opinions regarding the scientific basis for arguments found in the plaintiffs' motion for class certification. According to that motion (June 30, 2020) the plaintiffs have defined a "Minors Subclass" as follows:

> "All children who, during the period from May 1, 2014 to January 5, 2016, were (a) in utero or between the ages of 0 to 10 years old, (b) lived in an identified residence or attended an identified school or day care, and (c) were exposed through ingestion to unfiltered Flint public water at such residence school or day care for at least 14 days within a 90-day period."

In an accompanying footnote, the plaintiffs write that to be "exposed through ingestion to unfiltered public water" means that the children were exposed to unfiltered tap water, including infant formula, and/or ate food prepared with unfiltered Flint tap water or, alternatively, that the mother was similarly exposed.

It is the plaintiffs' claim that each individual in the "Minors Subclass" was harmed by the exposure as defined above and that that harm is defined, at least in part, in terms of neurodevelopmental outcomes, including intelligence (i.e. a decrease in IQ) as well as behavior (e.g. inattention, hyperactivity, antisocial behavior, among others). For confirmation, see the plaintiffs' experts' declarations by Drs. Lanphear and Hu. Although I will have more to say about these experts' declarations later in this report, it is important to put the plaintiffs' experts' arguments and claims in perspective because they relate directly to the claim that the "Minors Subclass" should be certified as a class.

To set the stage, I describe some plaintiffs' experts' arguments regarding the proposed subclass. For example, Dr. Hu writes the following [Declaration, p. 9-10 @ (3)]:

> "Each child who meets the criteria proposed in the subclass definition, and as further set forth herein, the criteria on which the definition is based, will more likely than not have experienced lead exposure as a result of the Flint water crisis. It is my opinion that the exposure is of sufficient duration and magnitude such that each child will have sustained a non-negligible impairment of their neurobehavioral development."

Dr. Hu continues by noting that evidence of the exposure to each member of the subclass (as defined above) is "fragmentary" and "sporadic" and therefore will not be considered in determining the extent to which the subclass was exposed to lead by ingestion during the Flint Water Switch. See Hu Declaration (p. 12, @ 10). Furthermore, he writes that even if blood lead levels were available for each member of the subclass, those measurements "cannot inform whether the lead came from absorption of lead from tap water vs. lead that had come from other sources (e.g. soil, paint, internal stores of lead in bone)." As a result, Dr. Hu (and selected other class certification plaintiffs' experts) rely solely on hypothetical measures of lead exposure for the members of the subclass.

Dr. Hu also writes that it is not necessary to know the IQ of the children—as defined in the "Minors Subclass"—before and after the Flint Water Crisis to make claims about the impact of hypothetical lead exposure on the intelligence (i.e., IQ) of those same children. He believes that actual measurements of IQ are, in his view, imprecise and error prone. He maintains this belief despite the fact that the studies that he and the other plaintiffs' experts rely on—e.g., the studies that were pooled to create the Lanphear et al. (2005) analysis—measured blood lead levels and IQs of each study participant at the individual level.

In sum, Dr. Hu and other plaintiffs' experts claim that they know what has happened and what will happen to the entire subclass of children—as a class—without measuring how much lead they were exposed to (as individuals) or their individual blood lead levels, and without knowing their individual IQs or their individual behavioral disorder status. In essence, the plaintiffs' experts ignore individual characteristics of members of the "Minors Subclass" in making their claims.

Other plaintiffs' experts, namely Drs. Ducatman and Keating, claim that the "Minors Subclass" as a group should also be provided with a medical monitoring program that, as I will describe in more detail, does not define the specific tests involved, the specific interventions, or the ethical safeguards essential to programs of this type. My opinions on medical monitoring can be found in a later section. I continue with my opinions regarding the scientific legitimacy of class certification regarding lead and IQ and behavioral disorders.

Opinions Regarding the Scientific Legitimacy of Class Certification

The causal nature of childhood neurodevelopment outcomes is a complex, multidimensional issue. The constellation of symptoms, diagnoses, measures of cognitive and behavioral outcomes, exposure characteristics and causal explanations for those events at the individual level are unique to those individuals. While it is possible to measure the per cent of children with elevated blood lead levels in a community, the average IQ of a group of children or the per cent of children in a group diagnosed with attention deficit hyperactivity disorder in epidemiological studies, the causal explanation of a change in an IQ score or occurrence of ADHD in any individual member of that group is, fundamentally and unequivocally, unique to that individual.

The plaintiffs' experts would have the reader believe that the only factors that are relevant to the occurrence of these outcomes in the "Minors Subclass" are age, residing in Flint, Michigan, and drinking water during the Flint Water Switch. They do not even claim that the "unfiltered water" drunk by an individual child was known to have been contaminated with lead. They further argue that all those in the "Minors Subclass" were exposed to lead—based on a hypothetical simulation exercise--in sufficient concentration and duration during the Switch that each member of the group experienced a decline in IQ and was put at risk of a neurobehavioral disorder (e.g. ADHD). The plaintiffs' experts' arguments—which are developed more fully below in this report—are overly simplistic, scientifically invalid, and in the case of the medical monitoring program, also ethically inappropriate.

The complexity of the issues involved in this legal matter can be illustrated with lead exposure itself. It cannot be assumed that every individual member of the "Minors Subclass" was exposed to additional lead as a direct result of the Flint Water Switch. It cannot be assumed that every individual member of the "Minors Subclass" was exposed to a sufficient concentration and duration of lead in drinking water that resulted in their blood lead level being raised to any level, much less to a level above 5 µg/dL or, for

that matter, above 7.5 µg/dL. As I will describe in detail, the plaintiffs' experts make no claim that they know the blood lead level or the measured amount of lead in the drinking water for each of the individual members of the "Minors Subclass" before the Flint Water Switch or during the Flint Water Switch. Indeed, the plaintiffs' experts basically ignore the data on blood lead levels that is available for individual representative class plaintiffs—see Dr. Hu's statements above—and thus rely solely on hypothetical blood lead levels generated from a statistical model, a simulation model that has not been tested against empirical data from the residents of Flint, Michigan including the representative class plaintiffs.

The amount of lead exposure and thus the potential effects of that exposure for any individual in the "Minors Subclass" are individual characteristics and not common characteristics. Even if all individuals in the "Minors Subclass" drank unfiltered water during the times defined by the plaintiffs, the extent to which lead existed in that water, how much lead was in the water, how much water the individual child drank and for how long, and the nutritional state of that child are not common characteristics of the members of the subclass. Each of these would vary based on the individual characteristics of the home and the child.

The same can be said for the many environmental factors—broadly defined—other than lead that can cause declines in IQ or the occurrence of behavioral disorders, including ADHD. These factors include maternal characteristics [e.g., age, race, education, smoking, alcohol use, use of illicit drugs, medical conditions (including preterm birth and other obstetric complications), pharmaceuticals, nutrition, weight and weight gain (during pregnancy), maternal IQ, marital status, and income] each with its unique exposure-response relationship. In addition, environmental chemicals and physical factors—sometimes referred to as environmental "toxins" at certain exposure levels and at specific ages, also cause adverse effects on neurodevelopmental outcomes. These include metals other than lead (e.g. methylmercury and arsenic) and other chemicals such as phthalates, polyhalogenated organics, pesticides (e.g. chlorpyrifos and the organophosphate pesticides), and brominated flame retardants. Details can be found in Part Two of this report.

These factors, whether considered separately or in combination, are alternative explanations for the occurrence of neurodevelopmental outcomes in the individual members of the "Minors Subclass." Each of these alternative explanations would vary across the members of that subclass. It is inconceivable that all the children in the "Minors Subclass" had mothers of the same age, race, education, smoking status, alcohol use, use of illicit drugs, medical conditions, weight and weight gain (during pregnancy), maternal IQ, marital status, and income. It may be reasonable to assume that a few of these factors could be the same or very similar for some members of the subclass. But it is not reasonable to assume that these factors are the same or very similar across all members of the subclass. Yet that is precisely what the plaintiffs' experts assume. They do not take into account these maternal factors. The same can be said for the environmental chemical and physical factors listed above. It is unreasonable to assume that the extent to which each child in the "Minors Subclass" had been exposed to one much less all these many factors known to affect IQ and other neurodevelopmental outcomes was the same for all members of the subclass.

It follows that to carefully and appropriately determine if exposure to lead from the Flint Water Switch was responsible for any specific neurodevelopmental outcome measured in an individual member of the "Minors Subclass," the extent to which any one of these many environmental factors or some combination of factors has affected that outcome in that individual must first be established. In addition, the extent to which that same individual has been exposed to each of these factors must be

documented to determine whether any such factor could potentially be responsible for an alleged neurodevelopmental outcome in that same individual rather than lead. From a scientific perspective, these individual determinations must occur before it would be reasonable to consider whether the "Minors Subclass" as a group had been harmed, given that the harm—according to the plaintiffs' complaint—must be common to all members of the group.

I recognize that exposure to lead in drinking water during the so-called "Flint Water Crisis" is a primary concern. However, the appropriate scientific approach to take before lead in drinking water can be considered a potential explanation for an observed cognitive or behavioral symptom, diagnosis, or test result in an individual who resided in Flint, Michigan during the crisis, is to define all sources of lead (before, during, and after the "Crisis") at the individual level. In addition, information must be known on each alternative explanation—the "environmental factors" described above—and the details of their relationship to the symptoms, diagnoses, or neurodevelopment test results found in the individual. These include the individual's mother's age, race, educational level, smoking history, history of alcohol use, history of use of illicit drugs, medical conditions (including preterm birth and other obstetric complications), use of pharmaceuticals, nutrition (and extent of prenatal care), weight and weight gain, IQ, marital status, and income. In addition, the extent to which the same individual was exposed (before, during, and after) the "Flint Water Crisis" to metals, solvents, other chemicals, pesticides, and brominated flame retardants must also be defined.

Simply put, there are at least 20 factors other than exposure to lead in drinking water (or lead from other sources) that can cause (or increase the risk of) any single neurodevelopmental outcome observed in an individual whether they were known to be exposed to lead in drinking water as a result of what is sometimes referred to as the "Flint Water Crisis."

Importantly, for any individual, accurate blood lead levels before and after the so-called "crisis" must also be documented, just as information on the existence and levels of each of the relevant alternative factors—(whether causes or risk factors)—must be known for the same individual before and after the so-called "crisis." Finally, information on IQ and other outcomes before and after the Flint Water Switch must be known at the individual level. Only with this information can claims about the role of lead from the Flint Water Switch be scientifically justified in the occurrence of neurodevelopmental outcomes at the individual level. These requirements are especially relevant given that the plaintiffs believe they can predict IQ declines from blood lead levels using a curve generated in the Lanphear et al. (2005) publication to be described in detail later in this report. Making predictions from that curve requires that the starting points for the prediction—the lead levels—must be known. In other words, lead levels before and after the Flint Water Crisis must be known. The reason this information is so important will be discussed in more detail later in this report. In sum, two of the plaintiffs' experts, Drs. Lanphear and Hu, disagree on the shape of the curve generated in the Lanphear et al. (2005) publication. For Dr. Lanphear, the curve is a curve, but for Dr. Hu the "curve" is actually a straight line. Dr. Lanphear's version of this curve—the relationship between blood lead levels and IQ decline—requires that the starting point be known at the individual level because IQ decline varies by starting blood lead level. In Dr. Hu's version, the starting point is irrelevant. In the end, Dr. Lanphear's version is that which was published. Dr. Hu is incorrect.

Given these considerations described above, a claim that all children living in Flint, Michigan during the Flint Water Switch—as defined in the "Minors Subclass"—were causally affected in their neurodevelopment by exposure to additional lead in drinking water due to the water source switch is scientifically untenable in the absence of individual assessments. Rather, each individual child who the

plaintiffs' experts' claim to have been affected by this additional lead (as described above) requires a solution to a separate and unique scientific problem of individual (i.e. specific) causation. In short, individual issues outweigh the common issues.

My opinion assumes that general causation relationships have been established between exposure to lead at the levels experienced by the children in the "Minors Subclass" and cognitive decline as well as behavioral disorders. General causation is a requirement for proceeding with individual causation assessments. These relationships, however, as discussed in detail in this report, have not been established especially at low levels of exposure, i.e. below 5 µg/dL and between 5 µg/dL and 7.5 µg/dL.

For all the reasons found above and further delineated in this report, certification of the plaintiffs' "Minors Subclass" regarding neurodevelopmental outcomes cannot be justified on scientific grounds.

Summary

The arguments of the plaintiffs' experts regarding class certification fail for all the reasons described immediately above. In sum, their arguments are scientifically bankrupt because they fail to take into consideration the fact that exposure to lead, the effects of lead, the effects of other causes of declines in IQ and risk of behavioral disorders are extremely unlikely, from a scientific perspective, to be identical across all members of the "Minors Subclass." To be more specific, the failure of the plaintiffs' experts' case regarding class certification stems from the following:

1. Failure to account for the extent to which exposure to lead through drinking water—existence, duration, and concentration—and other sources of lead exposure occurred for individual members of the "Minors Subclass."
2. Failure to account for the real possibility that factors other than lead exposure—the many environmental factors described above—were responsible for IQ deficits or behavioral disorders in the members of the "Minors Subclass" and that these factors are likely present in and affect each member of the subclass differently.
3. Failure to account for the relevance of individual measured blood lead levels or IQ measurements of the members of the "Minors Subclass" and relying therefore solely on hypothetical data generated by an untested assumption-laden simulation modeling exercise. Note that blood lead levels are available for representative class plaintiffs; see Section 4.2 of this report.
4. Failure to resolve the relevant and important disagreement among two of the plaintiffs' experts, Drs. Lanphear and Hu on the issue of the shape of the relationship between blood lead levels and IQ decline.

Section 1.5 Executive Summary Part II: Additional Opinions

Causal associations between exposure to low levels of blood lead (≤ 5 µg/dL) and neurodevelopmental outcomes (whether cognitive or behavioral) in children have not been established given the absence of effective adjustment for well-established causes and risk factors—i.e. confounders—of these neurodevelopmental outcomes in published epidemiological studies, the problems with the reliability of tests for low blood lead levels, and sparse epidemiological data. Note that this opinion negatively affects the viability and legitimacy of the plaintiffs' class certification claim. See also Section 4.2 of this report where blood lead levels of the representative class plaintiffs are tallied. All known values of

blood lead levels in all representative class plaintiffs are below 5 µg/dL as described in Section 4.2 of this report.

Causal associations between exposure to blood lead levels between 5 µg/dL and 10 µg/dL have not been established given the absence of effective adjustment for well-established causes and risk factors—i.e. confounders—of these neurodevelopmental outcomes in published epidemiological studies. This opinion also negatively affects the viability and legitimacy of the plaintiffs' class certification claim.

The pooled analysis by Lanphear et al. (2005) upon which the plaintiffs' experts rely is designed in such a way that does not provide a sufficient justification for establishing a causal relationship between blood lead levels and IQ in children. The design failures of this pooled analysis reflect the design problems of the original studies involved, namely, the use of concurrent blood lead and IQ measurements as well as failure to control for many established neurodevelopmental risk factors. As a result, the use of this analysis to predict IQ declines in individual members of the "Minors Subclass" is too uncertain to be scientifically valid and reliable.

The prevalence of causal factors and risk factors for adverse neurodevelopmental outcomes other than lead in women in Flint, Michigan during the years 2008-2015 is sufficient to explain the occurrence of those outcomes in children born during these years who were subsequently evaluated (e.g. in the years 2014 to the present). For examples, see Section 2.7 of this report. Put another way, given that the Lanphear et al. (2005) analysis fails to control for many of these factors—including but not limited to exposure to methylmercury, phthalates, PBDEs, and PCBs—it should not be used to make claims about the health of the children of Flint, Michigan who may or may not have been exposed to lead in water, depending upon their individual circumstances and who may have been exposed to a variety of other factors, i.e. well-established risk factors for neurodevelopmental outcomes other than lead. Note the relevance to the plaintiffs' class certification claim. Environmental factors that can affect IQ are present in Flint, Michigan and exposure to them cannot be assumed to be equal or even similar across all members of the "Minors Subclass."

Measured blood lead levels in young children (≤ 5 years of age) residing in Flint, Michigan during the switch from drinking water provided by the Detroit Water Authority (DWA) to the Flint River (FRW) revealed weak temporary elevations in average BLLs consistent with random variation. The average increase in BLLs among the Flint children was approximately 0.11 µg/dL, an increase that—if the relationship between BLLs and IQ is assumed to be causal—is, at best, associated with changes in IQ so small as to be clinically insignificant and uninterpretable.

An individual's blood lead level is a result of exposure to all prior and current sources of lead in the environment over the lifespan of that individual. For children, sources of lead exposure include but are not limited to dust, soil, water, food, toys, and paint. The blood lead value measured in any child is also affected by the child's age, mouthing behaviors, socioeconomic situation (including the extent to which lead exists in the plumbing of the child's home), iron status, and ethnicity. The relationship between lead and IQ that the plaintiffs rely upon in this litigation is primarily limited to that found in Lanphear et al. (2005) which is based on blood lead levels. In other words, the plaintiffs make claims of injury to all members of the class based primarily on the Lanphear et al. (2005) flawed analysis. Note the importance of individual rather than group (common) characteristics and the relevance to the class certification claim of the plaintiffs.

The opinions of Drs. Lanphear and Hu regarding general causation—i.e. opinions regarding the extent to which the relationship between exposure to lead and neurodevelopmental outcomes is causal—are scientifically invalid and unreliable on methodologic grounds. These will be described in more detail in Part Six of this report. In short, Drs. Lanphear and Hu do not provide systematic assessments of the scientific literature, fail to document claims with peer-reviewed scientific publications, and, in the case of Dr. Hu, make incorrect claims about the nature of risk. In sum, the scientific quality of their reports is poor, lacking key components of a valid and reliable scientific assessment of the issues involved.

The opinions of Drs. Ducatman and Keating regarding a medical monitoring program for the class in this litigation are scientifically invalid and unreliable. In addition, these experts have failed to provide any assurance that ethical imperatives of such programs—e.g. informed consent and the balance of benefits and risks of early detection—have been considered much less included. The medical monitoring program proposed by these experts cannot be said to be based on evidence that the benefits outweigh the risks. No mention much less a description of the informed consent process is provided in their reports. As such, it would be scientifically and ethically inappropriate to apply the proposed medical monitoring program to any individual member of the "Minors Subclass" much less to the entire "Minors Subclass."

It has not been established that there is a causal relationship between exposure to low-levels of lead (e.g. < 10 µg/dL) and renal disease/renal dysfunction or between exposure to low-levels of lead and hypertension (increased blood pressure) in children. In addition, it has not been established that renal disease or hypertension diagnosed in adulthood can be causally linked to low-level exposure to lead in childhood. It has not been established that there is a causal relationship between exposure to low-levels of lead (e.g. < 10 µg/dL) as children and cardiovascular disease in adults. Similarly, it has not been established that there is a causal relationship between exposure to low-levels of lead as children and essential tremor in adults. See Appendices G, H, I, and J for details. It has not been established that low levels of lead (e.g. < 10 µg/dL) cause spontaneous abortions.

These opinions are made with a reasonable degree of scientific certainty.

Section 1.6 Executive Summary Part III: Failure of the Plaintiffs' Experts' General Causation Case

Due to the complexity of this litigation, it may prove helpful to outline the plaintiffs' experts' general causation case and provide critical comments on its validity and reliability. I will show that the plaintiffs' experts in this matter make deceptively simple but scientifically invalid and unsubstantiated claims regarding general causation. Their arguments fail for solid scientific reasons.

In essence, the plaintiffs' experts argue that there is no lower limit of exposure to lead that does not cause adverse neurodevelopmental effects, whether cognitive or behavioral or both. Their claim can be restated as follows: that *any* exposure to lead that results in *any* change in a blood lead level in children causes an adverse neurodevelopmental effect. As a result, they argue, every plaintiff who lived in Flint, Michigan during the water source switch was adversely affected. Why? Because the plaintiffs' experts believe every plaintiff—i.e. every member of the "Minors Subclass"—had some exposure to lead-tainted water caused by the water source in 2014-2015.[2] To them, it does not matter how much additional lead

[2] It is important to point out that the defendant, Veolia, in this litigation was involved in the Flint Water Switch for a relatively short period of time, from February 2, 2015 through March 12, 2015. None of the plaintiffs' experts address this issue, i.e. the potential for injury to the class during this time period.

was involved in the alleged exposure at the individual level, because *any* exposure results in a change in blood lead level and therefore causes an adverse effect, in particular, a decrease in IQ.

The alleged scientific support for the plaintiffs' experts' claims comes primarily from a single study, representing a pooled analysis of data from seven prospective epidemiological studies and authored by Lanphear et al. (2005, 2019). The investigators of this study produced a regression line—a curve. Importantly, that analysis involves only lead exposure and IQ. The Lanphear et al. (2005, 2019) publication has no information on exposure to lead and behavioral disorders (including but not limited to hyperactivity, inattention, and impulsivity). With that curve, the plaintiffs' experts' claim they can precisely predict what IQ deficit occurs in a given child given a measured blood lead level in that same child. Plaintiffs' experts extend (extrapolate) this curve down to zero lead exposure and use this extrapolation of this curve to "prove" that there is no threshold for the effects of lead on neurodevelopment. "No threshold" is another way of saying that *any* exposure to lead—regardless of how small-- that results in *any* change in blood lead level—regardless of how small a change—causes an adverse effect.

The regression curve from the Lanphear et al. (2005) study used by the plaintiffs' experts provides an estimate of how large (or small) the IQ change (i.e. deficit) will be given the change in blood lead level measured in a child (or group of children). Consider these specific examples: according to the arguments provided by the plaintiffs' experts, if a blood lead level in a child from Flint, Michigan changes from 2.0 µg/dL to 3.0 µg/dL during a 14 day period, then the IQ for that child will be reduced by exactly 0.908 points according to the Lanphear et al. (2005) curve. It is also true that if a blood lead level for a Flint, Michigan child changes from 2.0 µg/dL to 2.01 µg/dL, then the IQ for that child will be reduced by exactly 0.009 points. Both examples are completely consistent with the plaintiffs' experts' basic claims: that *any* change in a blood lead level in a child causes an adverse cognitive effect, i.e. there is "no threshold," and their claim that they can know (i.e. predict) exactly the magnitude of that effect. It is important to point out that neither of these scenarios can be rejected by the plaintiffs' experts. If they were to reject the miniscule IQ change from the miniscule change in blood lead level as described above, then they cannot claim that there is "no threshold." In addition, if they were to reject the miniscule IQ change scenario, then they must reject the curve itself as not providing a valid approach to predicting the effects of lead exposure. After all, the curve provides what appears to be very precise predictions of IQ change given a blood lead level change.

The simplicity of the plaintiffs' arguments could seem compelling. After all, their claims rely upon an epidemiological study that, in turn, provides a quantitative approach to determining the magnitude of the adverse effect on IQ. In addition, they believe that by repeating the phrase that "there is no threshold" for the effect of lead on IQ that statement somehow reinforces their claim. Note however, that the phrase about thresholds appears in a different form in governmental reports. Those reports state that "no threshold has been identified" which is not the same as "there is no threshold" and is also not the same as "no exposure to lead is safe." I will have much more to say about this issue later in this report. The bottom line is that the identification of a threshold is, to put it bluntly, impossible based on the current epidemiological literature. Assuming that no threshold exists in this situation, as the plaintiffs' experts do, is just bad—invalid and unreliable—scientific practice.

Simplicity, of the sort illustrated in the plaintiffs' experts' reports and testimony, is not the best measure of validity and reliability in science. Facts and evidence and adherence to valid methodology are what counts in science however complex those facts and methods are. Based on these important scientific maxims, it would be a mistake to accept the plaintiffs' arguments and claims. Why? Because the

plaintiffs' experts rely upon several invalid assumptions. Put another way, the plaintiffs' claims require blind acceptance of several false and unsupportable assumptions. In this report, I will show that these assumptions cannot be substantiated. As a result, the claims made by the plaintiffs regarding the causal effects of lead exposure on the plaintiffs are invalid and therefore unreliable. A brief explanation follows.

The Plaintiffs' Experts' False and Unsubstantiated Assumptions

Assumption #1: That the lead-IQ experience of the children in each of the 7 cohorts is best described as the regression curve Lanphear et al. (2005) calculated. Or, to put it another way, it is irrelevant to the plaintiffs' experts whether the data from each of the 7 cohorts sits exactly on the regression line produced by the authors of the pooled analysis. As a result, the plaintiffs' experts assume that the relationship between lead and IQ for the Flint, Michigan "cohort" of children is also exactly—"i.e. best"—described by that same regression curve.

Assumption #1 is clearly false. The data from none of the 7 cohorts falls near much less exactly on the pooled regression curve created by Lanphear et al. (2005, 2019). It follows that it is extremely unlikely that the Flint, Michigan cohort (to which the individual members of the "Minors Subclass" belong) also falls anywhere near the pooled regression curve. Furthermore, the data from one of the cohorts reveals a contradictory result: that IQ increases as lead exposure increases. (See Section 2.2.1 of this report).

Assumption #2: That analyses performed by the investigators of each of the 7 cohort studies in Lanphear et al. (2005) adjusted for all established risk factors and causes of IQ deficits other than lead. In other words, the plaintiffs' experts assume that the curve is unbiased and is, therefore, valid.

Assumption #2 is clearly false. None of the 7 cohorts adjusted for most of the well-established risk factors and causes of cognitive (i.e. IQ) deficits. The Lanphear et al. (2005) regression curve is more likely to be biased than unbiased. (See Section 2.2.1 of this report).

Assumption #3: That the study data that generated the curve in Lanphear et al. (2005, 2019) was of the following format: it involved measured changes in IQ at the individual level and measured changes in the blood lead levels at the individual level. By "individual level" I mean that each child involved in the 7 different studies had both changes in IQ and changes in blood lead levels measured and then analyzed. Simply put, if the curve is being used to predict the change in IQ given a change in blood lead level, then the data that created that curve is of that same format.

Assumption #3 is clearly false. As Lanphear et al. (2005) explicitly note, only a single IQ measure for each individual in the study and a single blood lead level were used in the analyses. However, the plaintiffs use that information to make claims about changes in IQ for the members of the class. (See Section 2.2.1 of this report).

Assumption #4: That they can predict the IQ deficit caused by a change in blood lead level when the blood lead level is very low (i.e. less than 5 µg/dL). Low blood lead levels are believed to be present in the individual members of the "Minors Subclass" but not actually measured during and immediately after the Flint water switch. The plaintiffs' experts assume they can predict IQ changes at very low lead levels because the curve has been extrapolated down from a level of somewhere below 5.0 µg/dL to the level of 0 µg/dL.

Assumption #4, if not false, is fraught with serious scientific uncertainty. Alternative models have been published (e.g. Crump et al. 2013) that describe a statistically relevant relationship between lead levels and IQ that effectively challenge the "no threshold" argument that lies at the center of the plaintiffs' case. Furthermore, the curve generated by Lanphear et al. (2005) at very-low blood levels is extrapolated, i.e. made without sufficient data, including no data as the "curve" approaches zero. (See Section 5.3 of this report).

Assumption #5: That very small changes in IQ (< 5 points) are clinically significant.

Assumption #5 is unsupported in the scientific literature. Very small changes in IQ can be explained by many factors, including major stressors in the child's life (e.g. divorce, illness), exposure to other neurotoxins (including those not adjusted for in Lanphear et al. (2005)), the standard population to which an IQ measure is compared—the "Flynn Effect"—and the normal variability associated with taking tests more than once. Furthermore, it is unreasonable to argue that if an individual's IQ on one day was measured at 100 and then measured several months or several years later at 98 that that difference means that same individual is now not as "smart" as they were before. There is an expectation of variability around IQ scores measured in the same individual at different times. (See Section 3.2.1 of this report).

Assumption #6: That predicting the IQ deficit in an individual Flint, Michigan plaintiff can be done without knowing the extent to which other causes of IQ deficits were present and then excluded for that same individual. Put another way, the plaintiffs assume that the only factor affecting the neurodevelopment of the members of the "Minors Subclass" was the lead from the switch in water sources during 2014-2015.

Assumption #6 is clearly false. Accepting it flies in the face of well-established methods for determining individual causation in the scientific literature on this topic (Cole, 1997; Reference Manual on Scientific Evidence, 2000). Excluding known alternative causes is a necessary step in any method for establishing individual causation. Furthermore, it can be shown that many of the known neurotoxins and other factors affecting neurodevelopment were and continue to be present in the environment of Flint, Michigan. (See Section 5.5 of this report).

Assumption #7: That a biological mechanism has been established (in animal studies and/or molecular studies) that shows that any lead exposure (however small) causes an adverse effect in IQ. Put another way, the plaintiffs' experts assume that a mechanism exists that explains not only the effects of lead on neurons or other cellular components but also explains the effects of lead on IQ, a poorly understood measure of intelligence.

After all, the IQ measured in an individual can only be validly interpreted in the context of the distribution of IQ measures in a standard population. Simply put, an individual's IQ level is dependent upon the population to which it is compared. IQ is not, in other words, a measure of some physical property, like inches or pounds. IQ is a complex test-based relativistic construct.

Assumption #7 is clearly false. The plaintiffs have not provided the mechanism as described. It is not acceptable, in this situation, to argue (as the plaintiffs' experts may) that providing a mechanism is not required or that a general mechanism—e.g. that lead generally and absent specific details affects the central nervous system—is sufficient to explain how lead affects IQ. As described above and in more detail in this report, the epidemiological evidence is weak and inconsistent and, most importantly, incapable of accurate prediction at low blood lead levels. The current method of causal inference in the scientific community places a high priority on having in place a biological mechanism especially when the epidemiological evidence is weak. No such mechanism has been proposed much less substantiated in the plaintiffs' experts' reports.

As I will show in more detail in this report, each of these assumptions is either false or unsupported in the scientific literature. Perhaps more importantly, the plaintiffs' experts have failed to establish the validity of any of these assumptions.

Summary

The plaintiffs' experts' simplistic arguments fail. Importantly, I will show in my report that the plaintiffs' experts—specifically, Drs. Lanphear and Hu—do not adequately address the problems I have identified. Their approach is to rely primarily if not solely on the Lanphear et al. (2005) analysis as if it represents some absolute "truth" that cannot be considered invalid. Their reliance on the Lanphear et al. (2005) and seven false and/or unsupported assumptions make their conclusions invalid and unreliable. Failure of the general causation claims of the plaintiffs' experts means that the plaintiffs claim about the need for class certification fails.

Section 1.7 Executive Summary Part IV: The Failure of the Plaintiffs' Experts to Provide an Evidence-Based Medical Monitoring Program with Early Detection at its Core

Here, I summarize the claims of Drs. Ducatman and Keating who together attempt unsuccessfully to provide an evidence-based medical monitoring program for the "Minors Subclass." The proposed medical monitoring program fails to satisfy the requirements Dr. Ducatman explicitly states for such programs. These requirements, as clearly stated in the ATSDR guidelines cited by Dr. Ducatman, are that the tests (assessments) and interventions find outcomes early and have been demonstrated to provide benefits that outweigh the risks of those assessments and interventions. These requirements are population-based, i.e. they would apply to all members of the class as a class. However, there is no mention much less discussion by these experts of how any of the components of their proposed medical monitoring program meet these requirements. In fact, Dr. Keating provides no information on interventions, writing only that these can potentially be found in "clearinghouses." There is no description much less discussion of the need for informed consent where the benefits and risks are communicated to the participants of the program. Informed consent is an ethical imperative of any early detection—medical monitoring—program. In addition, neither Dr. Ducatman nor Dr. Keating

provide any evidence that the benefits of their proposed program outweigh the risks of the program. It is an ethical imperative that the benefits of medical monitoring programs outweigh the risks.

The plaintiffs' experts' proposed medical monitoring program is inappropriate on scientific and ethical grounds.

PART TWO
LIFE COURSE EPIDEMIOLOGY: FACTORS AFFECTING NEURODEVELOPMENTAL OUTCOMES IN CHILDREN

Section 2.1 Introduction

Although this litigation has a focus upon the possible effects of lead on the neurodevelopment of children, it is important to begin with a discussion of the well-established discipline of "life course epidemiology" that studies the impact of factors (including but not limited to lead) on children's health. For the past twenty-five years, epidemiologists have studied the health of children using a conceptual and methodological framework called "life course epidemiology" (Fox, 1998; Ben-Shlomo and Kuh, 2002). Life course epidemiology is defined as "the study of long-term effects on later health or disease risk of physical or social exposures during gestation, childhood, adolescence, young adulthood, and later adult life (Kuh et al. 2003, p. 778).

Life course epidemiology integrates biological and social risk processes; investigators examine how these many factors so defined provide important clues to the etiology of diseases and health conditions. Of special concern to life course epidemiologists are exposures that occur during what are referred to as "critical periods," i.e. a "time window when exposures can have adverse effects on development" (Kuh et al., 2003, p. 780).

A good example of "life course epidemiology" involving critical periods of time involves the many social and chemical factors—broadly, "environmental factors"—that can occur during the prenatal and postnatal periods with effects on children's neurodevelopment, including cognitive and behavioral outcomes.

Section 2.2 Environmental Factors Affecting Neurodevelopmental Outcomes

Lead in drinking water is the primary exposure of concern in the Flint, Michigan (Veolia) litigation, with special emphasis on the effect of lead on neurodevelopment in children who may have been exposed as a result of the Flint Water Switch in 2014-5 during the prenatal, postnatal or other time periods. However, lead is not the only environmental factor that can affect neurodevelopmental outcomes in children, including outcomes such as IQ, academic performance, attention deficit hyperactivity disorder (ADHD), and other behavioral disorders including associated symptoms (e.g. inattention, hyperactivity, and impulsivity). Furthermore, drinking water is not the only source of children's exposure to lead. Other sources of lead exposure include dust, soil, food, toys, and paint.

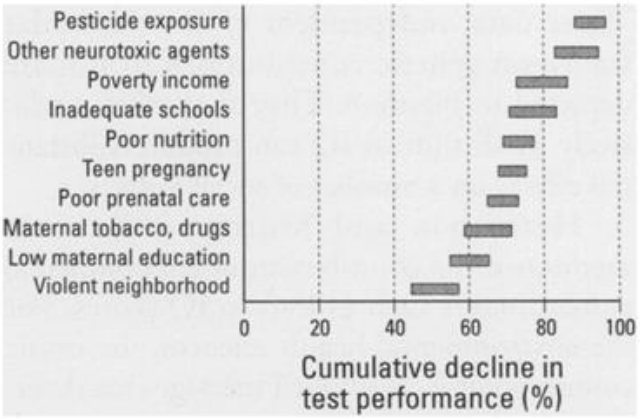
The list of factors known to adversely affect neurodevelopmental outcomes in children is long, especially if the phrase, "environmental factor" includes maternal characteristics such as age, race, education, smoking, alcohol and/or drug use, maternal weight or weight gain (e.g. overweight/obesity), maternal IQ, marital status, income, and other relevant demographic and socioeconomic indicators such as preterm (premature) birth (including late preterm and very preterm birth) and other obstetric complications (e.g. low birth weight, very low birth weight), nutrition, and prenatal care. Other important "environmental" factors are the social, emotional and cognitive support systems made available to the child in the home environment, captured by the Home Observation for Measurement of the Environment ("HOME") inventory (Bradley et al. 1988). The "HOME" inventory is commonly referred to and adjusted for in studies of blood lead levels and neurodevelopmental outcomes in children. However, the extent to which the many other factors that affect neurodevelopment have been adjusted for in the studies on lead is a matter of considerable concern and importance. In the

absence of such adjustments, it is difficult if not impossible to separate the effects of lead (if any) from the effects of these many factors.

The phrase "environmental factor" can also be defined narrowly, to mean specific chemicals and chemical mixtures found in the environment that also adversely affect neurodevelopment in children. The list of such "neurotoxicants" remains long enough to be a serious concern (Grandjean and Landrigan, 2014). Consider, for example, the chemicals other than lead that are known to affect children's neurodevelopment; these include methylmercury, polychlorinated biphenyls (PCBs), arsenic, toluene, manganese, fluoride, chlorpyrifos (and other organophosphate pesticides), dichlorophenyltrichloroethane, tetrachlorethylene, and polybrominated diphenyl ethers (PBDEs) (Grandjean and Landrigan, 2014). Some consider trace elements such as methylmercury, lead, and arsenic to be "paradigms" of developmental neurotoxicants (Grandjean and Herz, 2015). By that they mean that these compounds represent examples—sometimes called "exemplars"—of the role of neurotoxicants in children's health. Grandjean et al. (2017, p. 157) write that "lead, methylmercury, organophosphate pesticides, PCBs, PBDE, and air pollution are well documented causes of neurodevelopmental delay, deficits, and neuropsychiatric diagnoses."

It is important to keep in mind the fact that the "effects produced by a neurotoxic agent depend on many factors, including the timing and duration of exposure, the distribution of the toxic agent in different parts of the nervous system, the amount or concentration of the agent in nervous tissue, and the ability of the toxic agent to interfere with specific neurodevelopmental processes" (Jurewicz et al. 2013, p. 200-1). See also Rice and Barone (2000) for a detailed discussion of normal neurodevelopmental processes and critical periods of vulnerability for the effects of neurotoxic agents.

A legitimate scientific and public health goal is to understand how these many factors can impact children's neurodevelopment. In this matter, the scientific challenge is to determine, if possible, the potential effects of exposure to lead on children's neurodevelopmental outcomes independent of the effects of these other "environmental factors." As Weiss (2000) discusses in his review of the vulnerability of children and the developing brain to neurotoxic hazards, these many factors overlap in their effects in many different analytical forms. Weiss' conceptual model is shown immediately below, showing "how individual components of a stressful environment might cumulate to reduce performance on IQ and other tests. The individual stressors are shown as overlapping to suggest a lack of independence, and their length is meant to indicate that no single component is overwhelming in isolation" (Weiss, 2000, Fig. 5, p. 379).



Together, the environmental factors related to maternal characteristics (including demographic and socioeconomic indicators) and the factors encapsulated in the "HOME" inventory as well as the chemicals and chemical mixtures found in the environment—sometimes referred to as "environmental toxins"—make for a long list of causes and risk factors for adverse neurodevelopmental outcomes in children.

Indeed, Weiss, 2000 also evaluated the various confounding factors that are associated with cognitive deficits (e.g. pesticide exposure, other neurotoxic agents, poverty, inadequate schools, poor nutrition, teen pregnancy, poor prenatal care, maternal tobacco and drug use, low maternal education, and violent neighborhood) and noted that each factor contributes approximately 4% of the variance observed in IQ or roughly the equivalent of 10 ug/dL of exposure as measured via blood. Lead, on the other hand, contributes only approximately 1-6% of the variance. Put another way, somewhere between 99% and 94% of the observed neurodevelopmental effect in children can be explained by factors other than lead.

A review published in 2013 identified the following environmental toxins that in the authors' words present "strong and rather consistent indications that the developing nervous system is particularly vulnerable to insult from low levels of exposure to widespread environmental contaminants such as phthalates, bisphenol A, brominated flame retardants, polycyclic aromatic hydrocarbons (and) gas cooking" (Jurewicz et al., 2013, p. 185). A more recent commentary on contaminants found in drinking water that can negatively affect the developing brain includes the following: metals (e.g. lead, methylmercury, and arsenic), solvents, and industrial chemicals, including phthalates, polyhalogenated organic molecules, and bisphenol A (Silbergeld, 2016). In addition, Burnett et al. (2018, p. 485) write that "preterm birth (birth at <37 weeks of gestation) is a substantial risk factor for poor neurodevelopmental outcomes in areas such as neurosensory, cognitive, and behavioral functioning." Schug et al. (2015) of the National Institute of Environmental Health Sciences identified polychlorinated biphenyls (PCBs), phthalates, bisphenol A, metals, and pesticides as environmental chemicals associated with impaired neurodevelopmental outcomes.

Of special concern in this litigation is the extent to which these environmental factors (broadly defined) are responsible for effects on neurodevelopmental outcomes in those children who allegedly were exposed to lead in the drinking water in Flint, Michigan. When evaluating the validity and reliability of studies that purport to show the effect on neurodevelopment of exposure to lead in drinking water, it is critically important to assess the extent to which these "other contaminants," whether broadly or narrowly defined, were controlled (or adjusted) for in the statistical analyses or by restriction. In the absence of effective control (or adjustment), the specter of confounding raises its head reducing the validity of the inferences that can be drawn from a single study or from a collection of studies (Bellinger, 2004).

Put another way, to the extent that each of these factors affect neurodevelopmental outcomes— e.g. maternal factors, social factors, and chemicals and chemical mixtures—they can be considered alternative causes to any effect on neurocognitive or behavioral outcomes hypothesized to be due to lead exposure. Furthermore, these same factors—whether causal or not—can confound the hypothetical relationship between exposure to lead in drinking water and neurodevelopmental outcomes.

Before describing the theory behind the role of alternative causes and confounders in the study of lead and neurodevelopmental outcomes, it will be helpful to provide an illustration of how these many

environmental factors can affect causal inferences about the role of low-level lead and neurodevelopment when these factors are not measured or are poorly measured. A good example involves the pooled study of exposure to low levels of blood lead (BPb) and IQ as reported by Lanphear et al. (2005), re-analyzed by Crump et al. (2013), and prominently cited in influential reports such as the National Toxicology Program's Report on the Health Effects of Low-Level Lead (NTP, 2012) and which, according to Crump et al. (2013, p. 785) "has played a prominent role in shaping the public understanding of the effects upon children's IQ of low BPb exposures (e.g. BPb ≤ 10 µg/dL)."

Section 2.2.1 An Example: A Pooled Study of Prospective Cohorts Involving Children's IQ and Low-Level Lead Exposure (Lanphear et al. 2005)

Lanphear et al. (2005) is the report of a pooled analysis of the relationship between low-level environmental lead exposure and children's intellectual function using data from seven prospective cohorts: Boston (Bellinger et al. 1992), Cincinnati (Dietrich et al. 1993), Cleveland (Ernhart et al. 1989), Mexico City (Schnaas et al. 2000), Port Pirie, Australia (Baghurst et al. 1992), Rochester, New York (Canfield et al. 2003)), and (the former) Yugoslavia (Wasserman et al. 1997). The primary neurodevelopmental outcome measured in these studies was full-scale IQ, which includes verbal and performance scales. Several different versions of the Wechsler Intelligence Scale for Children were used in these studies including WISC-R, WISC-III, WISC-S, and WPPSI. IQ tests were given to children between the ages of 4 years and 10 years.

Data on a small group of potential confounders of the lead-IQ relationship were analyzed in these studies, including the child's gender, birth order, birth weight, maternal education, maternal age, marital status, prenatal alcohol or smoking exposure, as well as the HOME inventory. Only a single IQ measure and a single blood lead level were used from each of the 7 study datasets.

Importantly, the potential confounders not controlled for in these analyses included exposure to other known neurotoxic chemicals such as methylmercury, arsenic, polychlorinated biphenyls (PCBs), toluene, manganese, fluoride, chlorpyrifos (and other organophosphate pesticides), dichlorophenyltrichloroethane, tetrachloroethylene, polybrominated diphenyl ethers (PBDEs), air pollution [including its many individual components such as $PM_{2.5}$ and polyaromatic hydrocarbons (PAHs)], phthalates, bisphenol A, fluoride, and solvents. Also not measured and therefore not controlled for in these analyses were nutritional factors, maternal obesity, and the presence (or absence) of adequate prenatal care. The fact that these factors were not controlled/adjusted for in any of the analyses increases the likelihood that the regression curve is biased.

The authors of this pooled study investigated four measures of blood lead:

1. Concurrent blood lead level (that measured closest in time to the IQ test)
2. Maximum blood lead level (peak level at any time before the IQ test)
3. Average lifetime blood lead level
4. Early childhood blood lead level (collected between 6 and 24 months of age)

Statistical analysis involved fitting regression models to the data, including linear, quadratic, cubic, and spline. The results revealed an inverse relationship between blood lead and IQ with an estimated 6.9 point IQ decrement (95% CI: 4.2-9.4) associated with an increase in concurrent blood level from 2.4 to 10 µg/dL. In addition, the authors reported that for a given increase in blood lead, the lead associated

intellectual decrement for children with a maximum blood lead level < 7.5 µg/dL was significantly greater than that observed for those with a maximum blood lead level ≥ 7.5 µg/dL.

The basic results of the Lanphear et al. (2005, p. 898) analysis are shown in Fig. 2 and Fig. 3 from the publication, reproduced here:

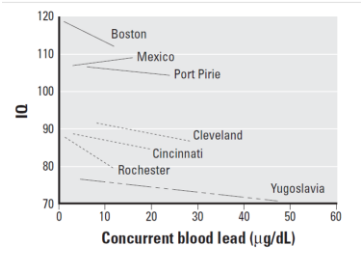


Figure 2. Linear models for each cohort study in the pooled analysis, adjusted for maternal IQ, HOME score, maternal education, and birth weight. The figure represents the 5th to 95th percentile of the concurrent blood lead level at the time of IQ testing.

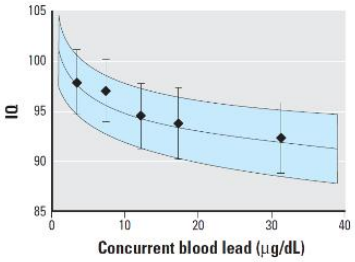


Figure 3. Log-linear model (95% CIs shaded) for concurrent blood lead concentration, adjusted for HOME score, maternal education, maternal IQ, and birth weight. The mean IQ (95% CI) for the intervals < 5 µg/dL, 5–10 µg/dL, 10–15 µg/dL, 15–20 µg/dL, and > 20 µg/dL are shown.

It is important to point out that the majority—perhaps even all—of the cohorts' data does not actually fall on the pooled regression curve. It follows that the curve—if it were measured—of the Flint, Michigan children potentially affected by the Flint Water Switch would also not fall on this pooled regression line. In scientific terms, there is high heterogeneity. Pooling data in this circumstance is inappropriate.

Lanphear et al. (2005, p. 894) concluded the following:

1. Environmental lead exposure in children who have maximum blood lead levels ≤ 7.5 µg/dL is associated with intellectual deficits
2. The observational design of the study limits (their) ability to draw causal inferences
3. The observed relationship—based on modeling—is a curve and not a straight line as shown immediately above in the Figure 3 from the Lanphear et al. (2005) publication

Nevertheless, the authors appear to believe their results, suggesting that those results support a causal connection between low-level blood lead and intellectual deficits in children.

Lanphear et al. (2005) briefly mention the possibility that confounding may have affected their results but they mention only three confounders: maternal depression, breast feeding, and iron status. They do

not mention the many confounding factors listed above that are established causes (or risk factors) of deficits in children's neurodevelopment. This is a serious deficiency of the Lanphear et al. (2005) analysis. As concerning is the fact that Lanphear et al. (2005) argue that maternal depression, breast-feeding and iron status are unlikely to be confounders in their data because these did not affect the results of earlier studies. This argument is scientifically invalid. The methodologic problem with this argument is that confounding is a feature of each individual study. Put another way, confounding must be evaluated in each study independent of the extent to which it did (or did not) affect the outcome measure in other studies. I will provide a more detailed account of the nature of confounding below.

Section 2.2.2 Adjustment for Neurotoxic Confounders in the Epidemiologic Studies of Low-Level Lead and Neurodevelopmental Outcomes

Later in this report, I will examine the extent to which other investigators attempted (or not) to control (i.e. adjust) for the many known causes and risk factors of neurodevelopmental outcomes in studies of lead exposure and neurodevelopmental outcomes. Details can be found below in Section 5. What this analysis reveals is that only a small fraction of the studies of lead and neurodevelopmental outcomes adjust for single factors much less for the full list of factors. It follows that the epidemiologic studies examining the hypothesis that low-level exposure to lead and neurodevelopmental outcomes in children have not provided valid tests of that hypothesis. In short, these studies have failed to exclude alternative explanations for the observed results, a critically important concern in scientific method. Failure to control for confounding—for such a long list of confounders—increases the likelihood that the curve is biased.

Section 2.3 Alternative Causes: A Primer and Application to the "Other Contaminants" Issue

Identification and exclusion of alternative causes—i.e. alternative explanations—is a key step in establishing general causation in the legal and scientific communities. This important concept as described in the Federal Reference Manual on Scientific Evidence (2000, p. 374) in the section entitled "Reference Guide on Epidemiology:"

> "In assessing causation, researchers first look for alternative explanations for the association, such as bias or confounding factors…"

See also (Reference Manual, 2000, p. 355):

> "…a study may reach incorrect conclusions about causation because although the agent and disease are associated, the agent is not a true causal factor. Rather, the agent may be associated with another agent that is the true causal factor, and this factor confounds the relationship being examined in the study."

It follows that all known causes and risk factors of an outcome—e.g. all known causes and risk factors of any specific neurodevelopmental outcome other than exposure to lead (i.e. the "other contaminants")—must be identified and excluded—through restriction or adjustment—before concluding that exposure to lead is responsible for (i.e. causes) an observed neurodevelopmental outcome.

Section 2.4 Confounding: A Primer and Application to the "Other Contaminants" Issue

Confounding is often defined as a mixing of effects between an exposure (e.g. exposure to lead), an outcome (e.g. IQ), and a third factor known as the confounder (e.g. any one of the "other contaminants" mentioned above). Confounding exists when the association between the exposure and the outcome is distorted by the impact of the confounding variable, specifically the relationships between the confounding variable and the exposure of interest as well as between the confounding variable and the outcome of interest (Aschengrau and Seage, 2003, p. 281-98). If, for example, an epidemiologic study of exposure to lead and IQ has failed to control for a contaminant that is associated with lead exposure and is associated with IQ, then it will be difficult if not impossible to ascertain the extent to which any observed "effect" is due to the lead exposure or due to the "other contaminant."

Any variable other than the exposure of interest can be a confounder if it has the following characteristics:

1. The confounding variable must be associated with the exposure in the study population, i.e. the confounding variable must be more or less common in the exposed groups than in the comparison (unexposed) group.
2. The confounding variable must be associated with the outcome of interest, either as a cause or risk factor.
3. The confounding variable cannot be an intermediate step in the hypothetical causal pathway between exposure and outcome of interest.

It is important to note that this description of confounding involves factors known to be associated with the exposure and the outcome. There are, in addition, factors unknown to the study investigators that can have the same impact as known confounders; these are referred to as "unknown confounders." Similarly, some known confounders may not have been measured in the study at issue or have been measured inaccurately. Observational studies—i.e. non-randomized epidemiological studies—cannot control for the unknown confounders, the unmeasured confounders, and the effects of poorly measured confounders.

Confounding effects—which diminish the validity and reliability of epidemiological studies—can be profound, especially in studies of lead exposure and IQ. Mink et al. (2004) reported that relatively small differences between exposed and unexposed groups in confounders such as maternal intelligence, home environment, and socioeconomic status (measured as years of parental education) could have clinically important effects on IQ by as much as 10 points. These authors note that their analyses were limited to just three factors. In other words, they did not examine how additional confounders (e.g. other known neurotoxic chemical contaminants) or other biases such as nonresponse, loss to follow-up, selection factors, missing data, or measurement error could alter the results of studies and decrease their validity.

The methodological issue of the confounding effects of "other contaminants" has been specifically discussed in several published papers regarding neurodevelopmental studies of children (Bellinger, 2004; Lanphear et al., 2008; Lanphear, 2015; Guth et al. 2020). Lanphear et al. (2005), for example, emphasizes the problem of the unmeasured confounders who write that "unmeasured confounding can be a particularly troublesome problem in observational epidemiology" (Lanphear et al. 2005, p. 197). The examples of unmeasured or poorly measured confounders given by Lanphear et al. (2005) include: exposure to pesticides, poor parenting, maternal depression, iron status, tobacco exposure, poverty, and pica. Lanphear et al. (2005) write that each of these represent potential confounders for the lead-diminished cognition relationship. However, any of the other factors described above—other metals

than lead, solvents and other industrial chemicals (phthalates, polyhalogenated organic molecules, and bisphenol A), parental alcohol and/or drug abuse, and maternal IQ—also meet the definition of a confounder. Guth et al. (2020) make perhaps the most important point of all: in order to argue that confounding has not occurred, it is necessary to have data, i.e. evidence of that fact.

Indeed, authors of studies of some well-established risk factors of neurodevelopmental outcomes in children (e.g. organophosphate pesticides) recognize that their results may be confounded by the presence of many other environmental factors. For example, Furlong et al. (2017, p. 13) write that "other environmental exposures, such as flame retardants, organochlorine pesticides or PCBs, air pollution, lead, heavy metals, as well as others, may confound our associations." The National Toxicology Program's report on the health effects of low-level lead (2012, p. 14) echoes this methodological maxim when they write that "in studies of outcomes causally linked to environmental tobacco smoke exposure, such as neurodevelopment or CVD (cardiovascular disease), environmental tobacco smoke may confound the observed association of lead and the health effect."

The solution to the problem of the unknown, unmeasured (or poorly measured) confounder is challenging. Lanphear et al. (2005, p. 198) argues that one solution would be to "rely on experimental studies of lead-exposed animals to confirm whether a metal or chemical is a toxicant." But this approach has serious methodological weaknesses. The identification of a specific human toxic chemical—that is, one that causes a specific effect in human populations—cannot be determined adequately by animal studies. Indeed, the premise of the Bradford Hill approach to causality (1965)—the so-called "Hill criteria"—is that there are nine considerations that require our attention, only one of which is the extent to which a chemical has been shown to affect mice, rats, hamsters, or other mammals, i.e. biologic plausibility. The rest of the Hill criteria have to do with well-defined characteristics of human epidemiological studies, whether the association exists, its strength, consistency, dose-response, specificity, and analogy with similar bodies of epidemiologic evidence, as well as temporality, experimentation and coherence. Bellinger (2004) makes a similar methodological error in causal inference in his discussion of the confounding problem. He writes that only two considerations are needed to make causal claims in the assessment of environmental neurotoxicant exposures and child neurobehavior: consistency (across epidemiologic studies) and evidence from animal model studies. Bellinger (2004) has excluded 7 of 9 of the Hill considerations, an invalid and unreliable inferential approach. Finally, neither quasi-solution offered by Lanphear et al. (2005) or by Bellinger (2004) eradicates much less reduces the fact that if known confounders are not adjusted for in epidemiological studies of lead and neurodevelopmental outcomes (e.g. due to lack of information or poor study design) then strong (i.e. valid and reliable) causal inferences regarding the hypothetical relationship between lead and the outcome(s) of interest cannot be established. This issue is particularly important when the lead exposure levels are very low (e.g. $<10\mu/m^3$) and the estimate of the hypothetical effect of lead on neurodevelopmental outcomes is relatively small (i.e. weak). It is a well-known phenomenon in causal inference that weak associations (i.e. weak "effects) are not good evidence of causality because they are so easily affected by confounding (whether known, unknown, or poorly measured).

One important issue remains regarding the characteristics of confounders in epidemiological research. Confounding is a study-specific feature. Put another way, the extent to which a factor confounds a result in one study population may not represent how that same factor confounds the result in a study of a different population. This occurs because confounding is dependent upon the extent to which the populations from which the exposed and unexposed groups are obtained in different studies can differ on the prevalence, distribution, and impact of each confounding variable. Confounding is a "property of

the source population" (Checkoway et al. 2004 p. 89). Each source population has unique characteristics. In practical terms, if a variable does not confound an observed exposure-outcome relationship in one study of lead exposure and neurodevelopmental outcomes, that does not mean that the same variable will not—or cannot—confound the same observed exposure-outcome relationship in a different study of lead exposure and neurodevelopmental outcomes. Each study population can and is likely to have different distributions of confounding variables.

Section 2.5 Confounders of the Lead-Neurodevelopment Relationship: Approach to their Identification

The approach to identifying confounders of the lead-neurodevelopment issue is to systematically search the English-language scientific literature for reviews, systematic reviews, and meta-analyses that have collected, described, summarized, and interpreted the scientific evidence on these topics. As mentioned above, there are several possible contenders for confounders of these relationships. The following were initially selected: alcohol, smoking, phthalates, obesity, mercury and methylmercury, polybrominated diphenyl ethers (PBDEs), polychlorinated biphenyls (PCBs), and pesticides. As described above, there are many other chemicals and chemical mixtures that could be included such as air pollution, arsenic, fluoride, and solvents. All are potential confounders of any hypothesized relationship between exposure to lead and neurodevelopmental outcomes.

Systematic Narrative Review, or "Evidence-based" Review

The systematic narrative review (sometimes called an "evidence-based" review) is a critically important methodology (Weed, 1997; Weed, 2000a, Weed, 2018). Here, the relevant scientific evidence is systematically collected, summarized, and interpreted. Systematic reviews are critical to the theory and practice of biomedical science. Systematic reviews not only summarize large bodies of information, but also inform medical and public health decision-making and guide policy and research priorities (1–7). Reviews enhance transparency and increase efficiency in the scientific process (8).

Although both narrative and systematic reviews will be collected and described here, systematic reviews are superior to narrative reviews in part because they are less prone to bias (Weed, 2018).

Systematic reviews begin with a research question. Then medical library databases are searched with a description of the search techniques made sufficiently transparent that the search could be repeated by others with similar if not the same results. In addition, it is common for authors of systematic reviews to supplement the searches with additional studies found in the reference lists of published papers or textbook chapters on the topic, government reports, and possibly, unpublished studies. The purpose of the review, the conditions (or criteria) for including and excluding the studies to be summarized and interpreted, and the criteria (or other methods) to be used in making causal claims are important—indeed, essential—components of a systematic narrative review. The word "narrative" is often used to partially describe this methodology, because it is common for authors to describe each study before the results are summarized and finally interpreted in terms of causation.

The discussion in the scientific community on the need for a systematic approach to reviewing the scientific and medical literature began in the mid-1980's and has continued unabated, with application to epidemiology as early as the mid-1990's (Milne and Chambers, 1993). For detailed discussions of the method and examples of its application, see: (Breslow et al., 1998; Crowther et al., 2007; Greenhalgh, 1997; Hutchison, 1993; Lichtenstein et al., 2008; Moher et al., 2008; Moher et al., 2009; Montori et al., 2003; Mulrow, 1987; Mulrow, 1994; Oxman et al., 1988; Oxman, 1994; Oxman et al., 2006; Rochon et

al., 2002; Shea et al., 2007; Weed, 1997; Weed et al., 2011, Alexander and Weed, 2016, and Weed, 2018).

Section 2.6 Summary

In Appendix A, systematic reviews were applied to the evidence on the relationships between the chemical environmental factors described earlier and adverse neurodevelopment outcomes. Based on the findings found in Appendix A, the following environmental factors are well-established causes (and/or risk factors) of children's neurodevelopmental outcomes: maternal smoking, maternal alcohol use, maternal obesity, preterm birth, exposure to methylmercury, phthalates, pesticides, air pollution, fluoride, polybrominated diphenyl ethers (PBDEs), polychlorinated biphenyls (PCBs), and exposure to cocaine and/or methamphetamines.

The table below summarizes the evidence described above regarding the extent to which maternal smoking, maternal alcohol use, maternal obesity, and prenatal exposure to phthalates, pesticides, air pollution, PBDEs (polybrominated diphenyl ethers), cocaine and methamphetamine increase the risk of neurodevelopmental outcomes in children. In this table, cognitive and behavioral outcomes are kept separate, recognizing that each general category involves a variety of outcome measures, and in some instances, gender-specific effects.

Exposure Factor	Cognitive	Behavioral
Maternal Smoking	Y	Y
Maternal Alcohol	Y	Y
Maternal Obesity	Y	Y
Preterm Birth	Y	Y
Methylmercury	Y	Y
Phthalates	Y	Y
Pesticides	Y	Y
Air Pollution	Y	Y
Fluoride	Y	?
PBDEs	Y	Y
PCBs	Y	Y
Cocaine	?	Y
Methamphetamine	?	Y

For additional confirmation of the effects of many of these same factors on neurodevelopment, see Bellinger (2012). There the author confirms that preterm birth, methylmercury, and organophosphate pesticides contribute greatly to the deficits in IQ in children. Note also that the author makes it clear that his goal "was not to provide an exhaustive accounting of all contributors (to IQ declines), and the selection of risk factors was not based on a systematic review…" (Bellinger, 2012, p. 501). Note also that in addition to the many established risk factors included in the table above, Bellinger adds several additional factors that impact IQ and could, therefore, be responsible for IQ deficits in the members of the "Minors Subclass." These additional factors include congenital heart disease, type 1 diabetes, acute lymphocytic leukemia, brain tumors, Duchenne muscular dystrophy, pediatric bipolar disorder, postnatal traumatic brain injury, failure to thrive, and iron deficiency.

In conclusion, the claims made by the plaintiffs' experts in this matter that IQ deficits in members of the "Minors Subclass" are necessarily caused by exposure to lead and not to other factors is an assumption that cannot be validly held. There are too many alternative factors—alternative causes—that have not been measured and controlled for to justify the plaintiffs' experts' claims.

Section 2.7 Environmental Factors other than Lead that Cause Neurodevelopmental Outcomes Prevalent in Flint City, Genesee County, and Michigan

In this section, the prevalence in Flint City of environmental factors other than lead that cause neurodevelopmental outcomes will be described. For some factors (e.g. PBDEs and PCBs) the available information is for Genesee County where Flint City is located. The importance of defining the prevalence of these factors in Flint City or Genesee County cannot be overstated; each of these factors represents an alternative explanation for any observed adverse neurodevelopmental outcome in children of Flint after the Flint Water Switch. I begin with a brief description of the location and demographics of Flint City and Genesee County, Michigan.

Flint City and Genesee County, Michigan

Flint City ("Flint") is the largest city in Genesee County, Michigan. According to the 2010 Census, there were 102,231 residents in Flint. The population has decreased by approximately 6% through 2018. In 2018, the U.S. Census estimated the population of Flint to be 95,932. Approximately 54% of the population in Flint is African American (n = 51,902) with 39% white (n = 37,069). The remaining ethnicities are relatively small groups of Native Americans, Asians, and Hispanics. In contrast, Genesee County had 425,790 residents in the 2010 Census and the county population has decreased to 406,892 with 20.3% African American and 75.4% white.

See: (www.census.gov/quickfacts/fact/table/flintcitymichigan,MI/IPE120218.

Approximately 40.4% of the Flint population live in poverty. In contrast, 18.8% of the population in Genesee County live in poverty, 14.1% of Michigan residents live in poverty and 11.8% of the U.S. population live in poverty. The 2018 per capita income in Flint was approximately $16,013 while the per capita income in Genesee County was approximately $26,386.

See: (www.census.gov/quickfacts/fact/table/Geneseecounty...).

Approximately 84.9% of Flint residents (> 25 years) have completed high school whereas only 12.0% have a bachelor's degree or higher. In contrast, 90.5% of Michigan residents have completed high school and 28.6% have a bachelor's degree or higher.

There are 83 counties in Michigan. For years, Genesee County has ranked relatively low on many health outcome and health factors according to a Robert Wood Johnson Foundation funded program with results found on the Genesee County Health Department website. The rankings are divided between two general categories: (1) health outcomes, including length of life and quality of life, and (2) health factors, including health behaviors, clinical care, social and economic factors, and physical environment. Genesee County rankings for health outcomes are shown below (relative to the other 82 Michigan counties with a rank of "1" as the best and "83" as the worst relative ranking). See: (https://gchd.us/resources/data).

Health Outcomes Ranking of Genesee County, Michigan

Factor	2011	2012	2013	2014	2015	2016	2017	2018
Health Outcomes	77	77	80	81	81	82	82	82
Length of Life	74	73	79	79	78	80	78	80
Quality of Life	79	77	81	81	81	81	82	81

Genesee County rankings for health factors are shown below, again relative to the other 82 Michigan counties.

Health Factors Ranking of Genesee County, Michigan

Factor	2011	2012	2013	2014	2015	2016	2017	2018
Health Factors	78	75	75	72	76	79	79	76
Health Behaviors	82	77	77	77	77	77	81	76
Clinical Care	28	18	18	16	22	20	21	27
Social + Economic	74	75	73	73	78	79	71	71
Physical Environment	23	67	76	78	75	79	80	82

Summary

With the single exception of clinical care, Genesee County ranks among the worst (bottom 10%) in the state of Michigan on health outcomes and health factors. Many of the more specific factors subsumed under these two broad categories are recognized as risk factors for adverse neurodevelopmental outcomes. Examples include low birth weight, smoking (adult), obesity (adult), excessive drinking, teen births, poverty, and single-parent households.

Prevalence in Flint, Michigan of Factors That Adversely Affect Neurodevelopmental Outcomes

As described earlier in this report, there are many environmental factors (defined broadly) that are known to adversely affect neurodevelopmental outcomes. These factors can be separated into two basic categories: (1) maternal and infant risk factors and (2) chemical and physical factors. It is important to describe the extent to which the residents of Flint, Michigan—and especially mothers and infants—have been and continue to be affected by these many factors that adversely affect neurodevelopmental outcomes.

Maternal and Infant Risk Factors Prevalent in Flint, Michigan (2008-2015)

Maternal risk factors for adverse neurodevelopmental outcomes include age (e.g. preteen pregnancies), low maternal education status, unmarried mothers, mothers in poverty, low maternal IQ, lack of prenatal care, smoking while pregnant, excessive alcohol use during pregnancy, and excessive weight gain (or obesity) during pregnancy. Infant characteristics include low birth weight (< 2500 g), very low birth weight (< 1500 g), preterm birth (< 37 weeks gestation), late preterm birth (34-36 weeks gestation), and very preterm birth (< 32 weeks gestation).

The extent to which the population of mothers and their infants in Flint, Michigan manifested these risk factors for adverse neurodevelopmental outcomes can be found in the following tables, adapted from the Michigan Department of Health website (www.mdch.state.mi.us) accessed April, 2020, for the years 2008 through 2015. These years were chosen because many of the youngest Flint children who may have been exposed to lead in drinking water as a result of the "switch" were conceived and born during

this time period. For example, a Flint child who was born in 2008 was 6 years old in 2014. Importantly, these prevalence values do not change significantly through 2018.

Prevalence of Maternal Risk Factors for Adverse Neurodevelopmental Outcomes: Flint (2008-2015)

Maternal Risk Factor	2008	2009	2010	2011	2012	2013	2014	2015
Under 20y[1] (%) "Teens"	19.9	22.7	19.0	17.3	16.4	15.2	14.5	12.1
< 12 y Education (%)	30.7	31.1	29.0	32.4	27.6	27.0	26.4	27.2
Unmarried (%)	78.0	80.2	79.0	82.3	81.4	83.5	83.8	83.1
No Prenatal Care[2] (%)	40.0	40.4	36.7	37.2	35.7	38.8	41.6	27.0
Smoked[3] (%)	28.4	28.6	31.3	29.7	30.4	30.9	29.3	28.1
Excess Weight Gain[4] (%)	50.9	52.2	51.6	50.3	51.1	51.9	51.4	51.2

[1] "Pregnant teens are more likely than older women to receive late or no prenatal care, have eclampsia, puerperal endometritis, systematic infections, low birth weight, preterm delivery, and severe neonatal conditions" (USDHHS, HSRA, Maternal and Child Health Bureau, Child Health, USA, 2011).

[2] Defined as no prenatal care in the first trimester.

[3] Defined as those mothers who did not quit smoking prior to pregnancy or chose to quit during pregnancy.

[4] Excess weight gain is defined in terms of the Institute of Medicine guidelines found on the American College of Obstetrics and Gynecology website (acog.org). For example, obese mothers (≥ 30 BMI) should gain no more than 11-20 pounds during pregnancy. Overweight mothers (25-29.9 BMI) should gain no more than 15-25 pounds during pregnancy. Women of normal weight (18.5-24.9 BMI) should gain no more than 25 to 35 pounds.

Summary

The prevalence of risk factors for adverse neurodevelopmental outcomes in women in Flint, Michigan during the years 2008-2015 is sufficient to explain the occurrence of those outcomes in children born during these years who were subsequently evaluated (e.g. in the years 2014 to the present). For example, approximately half of the Flint, Michigan pregnancies in the years 2008-2015 involved mothers who gained too much weight. Approximately 30% of the Flint, Michigan pregnancies involved mothers smoking. Close to 40% of the pregnancies in this same period involved no prenatal care during the critically important first trimester. Finally, between 15 and 20% of the Flint, Michigan pregnancies involved young (teen) mothers.

Prevalence of Infant Risk Factors for Adverse Neurodevelopmental Outcomes: Flint (2008-2015)

Infant Risk Factor	2008	2009	2010	2011	2012	2013	2014	2015
Low Birth Weight (%)	13.9	14.2	12.3	14.5	14.5	12.8	13.1	15.2
Very Low Birth Weight (%)	2.7	3.5	2.7	2.6	2.5	2.0	2.0	3.3
Very Preterm[1] (<32 wk) (%)	3.0	3.4	3.2	3.2	2.9	2.8	2.1	3.5
Late Preterm[1] (34-6 wk) (%)	9.7	10.0	9.1	10.3	10.7	9.1	12.0	10.4
Preterm[1] (<37 wk) (%)	14.7	15.5	14.0	15.4	15.3	14.2	15.7	15.6

[1] Prematurity of any degree affects the cognitive performance of children born preterm. The poor neurodevelopment persists at various ages of follow-up" (Allotey et al. 2018). "Preterm children born in the antenatal steroids and surfactant era (i.e. between 1990 and the present) show considerable academic difficulties" (Twilhaar et al. 2018).

Summary

The prevalence of infant risk factors (e.g. approximately 15% of births were preterm and 15% were of low birth weight) in Flint, Michigan births during the years 2008-2015 is an important source of adverse neurodevelopmental outcomes in children evaluated in later years (i.e. 2014 through the present).

<u>Prevalence of Chemical and Physical Factors in Flint, Michigan Affecting Neurodevelopmental Outcomes</u>

As described earlier in this report, there are many environmental factors (defined broadly) that are known to adversely affect neurodevelopmental outcomes. Here, I focus on chemical and physical factors and the extent to which the residents of Flint, Michigan—and especially mothers and infants—have been exposed to and continue to be affected by these many factors that adversely affect neurodevelopmental outcomes. Representative examples include: mercury and methylmercury, phthalates, polybrominated diphenyl ethers (PBDEs), and polychlorinated biphenyls (PCBs).

Background: Mercury (Hg) and Methylmercury (MeHg)

Mercury (Hg) is a naturally occurring heavy metal often found in rock and in coal deposits. On the periodic table, it has the symbol "Hg," and its atomic number is 80. Mercury has three basic forms: elemental (metallic) mercury, methylmercury and other organic compounds, and inorganic mercury (www.epa.gov/mercury).

Elemental or metallic mercury is a shiny silver-white liquid at room temperature. As a liquid, mercury is often used in older thermometers, fluorescent light bulbs and some electrical switches. When exposed at room temperature, elemental mercury can evaporate to become an invisible, odorless toxic vapor.

Mercury (elemental and/or inorganic) gets into the air from several sources including industrial waste (e.g. coal combustion), waste incineration, and mining activities (Ye et al. 2017). Once in the air, mercury settles into bodies of water or gets washed there from soil deposits (Castoldi et al. 2008). In turn, methylmercury is formed by the action of microbes that live in aquatic systems including lakes, rivers, wetlands, sediments, soils, and the open ocean.

The most common route of exposure to methylmercury in the United States is by eating fish or shellfish containing methylmercury; local sport fish as well as canned fish (e.g. tuna) are important sources of mercury (i.e. methylmercury) exposure. The most common route of exposure to inorganic mercury (e.g. cinnabar) is from occupational sources, including welding and mining. Exposure to elemental mercury can occur when thermometers break and in some novelty jewelry products as well as dental fillings (U.S. EPA website, accessed 4.2020). Exposure to mercury can also occur through inhalation (Ye et al. 2017).

Mercury Exposure in Genesee County, including Flint City

Mercury contamination of rivers and sport fish in Genesee County, Michigan is common. The Michigan Department of Health and Human Services produces a 2018 document entitled “Eat Safe Fish Guide” that describes edible sport fish species in specific Michigan locations with known methylmercury levels high enough to warrant limiting consumption. The fish species in Genesee County locations that have high levels of mercury are shown in the following table, adapted from the “Eat Safe Fish Guide” (2018, p. 24-27) at www.michigan.gov/eatsafefish:

Location	Black Crappie	Catfish	White Crappie	Largemouth Bass	Smallmouth Bass
Flint River	•	•	•		
Holloway Reservoir	•	•	•		
Kearsely Creek				•	•
Kearsley Reservoir				•	•
Lake Fenton				•	•
Lobdell Lake				•	•

Mott Reservoir	•	•	•		

Other sport fish species in Michigan have high levels of methylmercury, including walleye and northern pike.

Summary

Consumption of sport fish caught in Michigan, including fish from the Flint River and other locations within and near Genesee County, carries with it a risk of methylmercury exposure. Mercury exposure can also occur by consuming canned fish (e.g. tuna fish)

Background: Phthalates

Phthalates belong to a family of chemicals derived from phthalic acid and used in a variety of products including children's toys and medical devices primarily to make them soft and flexible (Jurewicz and Hanke, 2011; Jurewicz et al. 2013; Zhang et al. 2019). Phthalates are also added to cosmetics where they act as a vehicle for fragrance and to other personal care products. Phthalates are ubiquitous (Radke et al. 2020); they are found in adhesives, glues, building materials, detergents, paints, pharmaceuticals (primarily as enteric coating), textiles, floor tiles, cables and wires, and food wrappers and containers (Jurewicz and Hanke, 2011; Jurewicz et al. 2013; Ejaredar et al. 2015; Zhang et al. 2019). The most common phthalate used is one of the high molecular weight (HMWP) variety, namely di-(2-ethylhexyl)phthalate (DEHP) (Lee et al. 2018; Zhang et al. 2019). Other commonly used phthalates are diisodecylphthalate (DIDP), diisonoylphthalate (DINP) and benzylbutylphthalate (BBP). Exposure to phthalates can occur through many routes including oral (food, water, beverages, and teething), inhalation (hair spray, nail polish), parenteral (tubing in neonatal intensive care units), as well as by crossing the placenta (Jurewicz and Hanke, 2011). Phthalates do not bioaccumulate in the human body; their metabolites can be measured in blood, breast milk, and meconium (Jurewicz and Hanke, 2011; Ejaredar et al. 2015). Phthalates are sometimes referred to as endocrine disruptors because they can either mimic or interfere with human hormones (i.e. the endocrine system). They have been studies primarily for their potential impact on reproduction and the thyroid although studies have also been performed examining the impact of phthalates on neurodevelopment in infants and children (Ejaredar et al. 2015; Lee et al. 2018).

Phthalate Exposure in the U.S. Population

Over 75% of Americans have measurable phthalates in their urine (Lee et al. 2018). Exposure to phthalates occurs via ingestion, inhalation, and dermal absorption (Zota et al. 2014; Varshavsky et al. 2018). Studies of the extent to which the U.S. population has been exposed to phthalates often employ the data found in the National Health and Examination Survey (NHANES), a nationally representative survey of the civilian non-institutionalized population. In the decade 2001-2010 exposure to some phthalates have decreased while others have increased such that the overall exposure has remained relatively constant (Zota et al. 2014; Reyes and Price, 2018). Children tend to have higher levels of phthalates than adolescents or adults (Reyes and Price, 2018; Zota et al. 2014).

Phthalate Exposure in Michigan Residents

Studies revealing prominent examples of phthalate exposure in Michigan residents have been reported (Watkins et al. 2016; Goodrich et al. 2019). Researchers at the University of Michigan have measured

maternal phthalate exposure in women participating in the Michigan Mother-Infant Pairs cohort. Urine and blood samples were collected during the first trimester of pregnancy and at delivery. The authors note that exposure to phthalates during pregnancy is "of particular concern because hormonal disruption during fetal development may influence fetal growth with long-term effects on health" (Watkins et al. 2016, p. 2). The authors found that maternal DEHP metabolite concentrations were significantly higher at delivery compared to the first trimester and that a number of phthalate metabolites were associated with birth size and gestational age in patterns that varied by sex and timing of exposure. Simply put, there is a distribution of phthalate exposure across pregnant women in Michigan due in part to the fact that exposure to phthalates is multidimensional and depends on use of materials, diet, and other environmental conditions.

Polybrominated diphenyl ethers (PBDEs)

Polybrominated diphenyl ethers (PBDEs) are brominated hydrocarbons used as flame retardants in many consumer products and building supplies (US EPA, 2017). Examples include plastics, textiles, foams, electronic equipment, and upholstery (Batterman et al. 2009; Batterman et al. 2010). PBDEs are considered ubiquitous pollutants in the environment, are persistent and bioaccumulative (US EPA, 2017; Bradley et al. 2011). They can be found in houses, workplaces, and vehicle interiors with higher concentrations indoors than outdoors (Batterman et al. 2010). Exposure to PBDE can occur from vapor, particulate matter, or in floor dust. In addition, PBDEs can bioaccumulate in the food web and make their way to humans through fish consumption, drinking water, and recreational activities (Venier et al. 2014).

Polybrominated diphenyl ethers (PBDEs) in Michigan Residents and Michigan Waters

The North American Great Lakes, including Lake Michigan and Lake Huron are contaminated with PBDEs (Zhu et al. 2005, Bradley et al. 2011, Song et al. 2005; Venier et al. 2014). Sport fish contaminated with PBDE have been found in the Detroit River (Michigan) and in the Great Lakes (Rice et al. 2002; Zhu and Hites, 2004). PBDEs have also been found in office buildings in Michigan primarily in settled dust and indoor air (Batterman et al. 2010). PBDE exposure is common in the U.S. population (Yuan et al. 2017). Importantly, African American women living in Detroit, Michigan have had PBDE plasma levels investigated. A high proportion of these women had detectable levels of PBDEs in plasma. The most consistent indicators of elevated PBDE plasma concentrations were lower BMI, lower educational attainment, and heavy ethanol consumption. PBDE plasma concentrations were inversely correlated with age (Orta et al. 2020).

Background: Polychlorinated Biphenyls (PCBs)

https://www.epa.gov/pcbs/learn-about-polychlorinated-biphenyls-pcbs

PCBs belong to the family of man-made organic chemicals known as chlorinated hydrocarbons. PCBs were domestically manufactured from 1929 until manufacturing was banned in 1979 (US EPA, 2020; Schug et al. 2015). PCBs vary in consistency from thin, light-colored liquids to yellow or black waxy solids. Due to their non-flammability, chemical stability, high boiling point and electrical insulating properties, PCBs were used in industrial and commercial applications including electrical, heat transfer and hydraulic equipment, plasticizers in paints, plastics and rubber products, pigments and dyes and flame retardants (US EPA, 2020; Schug et al. 2015).

Typically, the PCBs used in commercial and industrial applications are chemical mixtures made up of a variety of individual chlorinated biphenyl components known as congeners. Most commercial PCB mixtures are known in the United States by their industrial trade names. Arochlor is one of the most common PCBs.

PCBs are ubiquitous in the environment and do not readily break down because of their resistance to degradation. They are a good examples of persistent organic pollutants (POPs). They can remain for long periods cycling between air, water, and soil. PCBs have been found in snow and sea water in areas far from where they were released into the environment. PCBs also accumulate in plants and food crops given their capacity for transport from poorly maintained hazardous waste sites, leaks from electrical transformers, disposal of PCB-containing consumer products into landfills, and burns (US EPA, 2020; Caspersen et al. 2016).

Furthermore, PCBs can accumulate in the leaves and above-ground parts of plants and food crops. They are also taken up into the bodies of small organisms and fish. As a result, people who ingest fish may be exposed to PCBs that have bioaccumulated in the fish they are ingesting.

PCBs have lipophilic properties (i.e. they accumulate in fatty tissue) and can cross the placenta. The extent of exposure of infants to PCBs prenatally depends primarily upon the mother's long-term exposure to PCBs from dietary sources (Caspersen et al. 2016; Vrijheid et al. 2016).

PCB Levels in the United States

Biomonitoring of the PCB levels in the U.S. population based on the National Health and Nutrition Examination Survey (NHANES) reveals the following median concentrations in blood serum, by race/ethnicity and family income, for the years 2001-2004, adapted from the report "America's Children and the Environment, 3rd ed found on the US EPA website, www.epa.gov:

Race/Ethnicity	All Incomes	< Poverty Level	≥ Poverty Level
All Races/Ethnicities	30.1[1,2]	25.8	31.8
White, non-Hispanic	33.6	29.0	34.8
Black, non-Hispanic	32.2	30.3	37.4
Mexican-American	18.0	16.1	18.9
Other Races/Ethnicities	31.6	NA	38.0

[1]Median concentration of PCBs in serum (ng/g lipid)
[2]Adjusted for the likelihood of getting pregnant

PCBs and Exposure in Michigan Residents

Polychlorinated biphenyls (PCBs) are prevalent in many fish species found in Lake Erie, Lake Michigan, and Lake Superior (Salamova et al. 2013; Rasmussen et al. 2014; Han et al. 2016). Dietary consumption of sport fish (from the lakes in and around Michigan) provides an important source of PCBs (Han et al. 2016). PCBs are, as discussed elsewhere in this report, are potent neurotoxicants causing deficits in cognitive function, including effects on IQ, memory, verbal abilities, motor skills, and visual-spatial skills given prenatal and early-life exposures (Turyk et al. 2012).

Summary

The prevalence of chemical and physical risk factors for adverse neurodevelopmental outcomes in women in Flint, Michigan during the years 2008-2015 is sufficient to explain the occurrence of those outcomes in children born during these years who were subsequently evaluated (e.g. in the years 2014 to the present). For example, methylmercury, phthalates, PBDEs, and PCBs are prevalent in the atmosphere and food web of residents of Michigan, Genesee County, and Flint City.

PART THREE: BACKGROUND ON SOURCES OF LEAD, COGNITIVE AND BEHAVIORAL DISORDERS

Section 3.1 Sources of Lead and Measurement

Lead is a soft, malleable, and dense blue-gray metal and ubiquitous in the environment. The level of exposure to lead for any given individual is dependent upon several factors, including occupation, geography, social circumstances, and stage of life (NTP, 2012; USEPA, 2020 Website). Lead can be found in soil, air, drinking water, consumer products, food, occupational settings, paint, ceramics, pipes and plumbing materials, solder, gasoline (in the past), batteries, ammunition, and cosmetics (USEPA, 2020, Website). Laidlaw et al. (2016, p. 1) note that despite regulations in the latter part of the 20th century that reduced the lead content of gasoline, paints, as well as other sources, the "damage was already done" inasmuch as the "insolubility of lead resulted in the accumulation and concentration of lead in surface soils and dust." Primary routes of lead exposure for children include oral exposure from hand to mouth behavior, especially in young children, from dust, toys, and paint. Other sources of oral exposure include contaminated food and water. Dietary sources of lead include imported foods (e.g. from Mexico), spices, herbs, and nutritional supplements, as well as traditional (folk) medicines. Infants can be exposed to lead through contaminated water used to make infant formula. Pottery can be covered with a lead-contained glaze that will leach into food or water placed in the pot. Lead crystal glassware leaches lead into alcohol-based drinks (Buettner et al. 2009; NTP, 2012; USEPA, 2020 Website).

The National Toxicology Program's Report on the Health Effects of Low-Lead Exposure (2012, p. 12-13) notes:

> "Tap water once contributed to as much as 10-20% of total lead exposure in the United States before amendments to the Clean Water Act and some older pipes, taps, and pre-1986 pipe solder still contain lead."

This same NTP publication describes how lead can enter drinking water through corrosion of lead from pipes and plumbing fixtures and after changes in disinfection processes, particularly with the use of chloramine citing the events in Washington, D.C. and described by Edwards et al. (2009).

Lead can be inhaled, for example, from dust created when renovating older buildings that used lead-based paints, especially those built before 1978. According to the NTP (2012) Report, the Department of Housing and Urban Development estimated that 40% of U.S. housing contained lead-based paint in 2001. Lead in soil and dust originated, at least in part, from extensive use of leaded gasoline (NTP, 2012, p. 13). Tobacco smoke also contains lead.

Children's exposure to lead is of special concern. The NTP (2012, p. 14) Report notes the following:

> "Children are most commonly exposed to lead in paint, household dust, and soil—particularly if they reside in pre-1978 deteriorated housing—and can increase their risk of exposure by natural mouthing tendencies."

See also Lanphear et al. (1998) and USEPA (2006). Lanphear et al. (2002, p. 45) also note that "children who live in rental housing (are) also at risk for higher blood lead concentration." Finally, children can also be exposed to lead brought into the home on the work clothes of adults whose work involves lead (CDC, 2009, p. 213).

Lead Blood Levels

Lanphear et al. (1998, p. 51) write that "…a child's age, race, mouthing behaviors, and study-site specific factors influence the predicted blood lead levels at a given level of exposure." Indeed, an individual's blood lead level reflects the equilibrium between environmental sources of lead and the internal body burden of lead (NTP, 2012). The half-life of lead in blood is approximately 1 month.

Measurement of blood lead at levels ≤ 5 µg/dL is particularly challenging as discussed by Caldwell et al. (2017). As these authors note (Caldwell et al. 2017, p. 1):

> "Review of five years of results for target blood lead values < 11 µg/dL for U.S. clinical laboratories participating in CDC's voluntary Lead and Multi-Element Proficiency (LAMP quality assurance program showed 40% unable to quantify and reported a non-detectable result at a target blood lead value of 1.48 µg/dL compared (with) 5.5% at a target blood lead of 4.60 µg/dL."

It follows that the probability of an accurate blood lead level measured below 5 µg/dL is less than that of a blood lead level measured above 5 µg/dL.

Blood lead levels in the U.S. population are monitored, at least in part, by referring to published studies that report data from the National Health and Nutrition Examination Survey (NHANES). NHANES is an ongoing cross-sectional survey study collecting and analyzing data on health and nutrition maintained by the National Center for Health Statistics (NCHS) of the CDC (www.cdc.gov/nchs/nhanes.htm). The NHANES study is designed to be representative of the U.S. national noninstitutionalized population and involves both interviews and medical examinations, including blood draws. Examples of the use of NHANES to monitor blood lead levels in children are briefly described in the following section. One way to evaluate the clinical and public health significance of these data is to compare the blood lead values reported in these NHANES publications with the "level of concern" or "reference value" promulgated by the CDC. "Levels of concern" and "reference values" change over time. A brief description of the CDC's decisions to create "levels of concern" and, more recently, "reference values" for blood lead levels in children begins this next section.

The CDC's "Levels of Concern" and "Reference Values" for Blood Lead Levels in Children

In a 2005 report (p.2) on the prevention of lead poisoning in children, the CDC stated the following:

> "Between 1960 and 1990 the blood lead level for individual intervention in children was lowered from 60 µg/dL to 25 µg/dL. In 1991 the CDC recommended lowering the level for individual intervention to 15 µg/dL and implementing communitywide primary lead poisoning prevention activities in areas where many children have BLLs >10 µg/dL. Some activities, such as taking an environmental history, educating parents about lead, and conducting follow-up blood lead monitoring were suggested for children with BLLs of >10 µg/dL. However, this level, which was originally intended to trigger communitywide prevention activities, has been misinterpreted frequently as a definitive toxicologic threshold."

Since 2012, the CDC has eliminated the phrase "level of concern" and "now uses a blood lead reference value of 5 micrograms per deciliter to identify children with blood lead levels that are much higher than most children's levels. This new level is based on the U.S. population of children ages 1-5 years who are in the highest 2.5% of children when tested for lead in their blood. This reference value is based on the 97.5th percentile of the National Health and Nutrition Examination Survey (NHANES)'s blood lead distribution in children. The current reference value is based on NHANES data from 2007-2008 and 2009-2010" (www.cdc.gov/nceh/lead/prevention/blood-lead-levels.htm).

Blood lead levels in U.S. Children based on the NHANES Data

Mahaffey et al. (1982) reported data on blood lead levels from the 2nd NHANES survey, conducted between 1976 and 1980. Average (mean) and median blood lead levels by age (all races) are shown in the following table for individuals ages 0.5 years to 17 years, adapted from Mahaffey et al. (1982, Table 1, p. 575):

Age Group (years)	# Persons Examined	Mean Blood Lead Level[1]	Median Blood Lead Level[1]
0.5-2	759	16.3±0.57	15.0
3-5	1613	15.9±0.40	15.0
6-8	451	13.9±0.47	13.0
9-11	377	12.9±0.39	12.0
12-14	448	11.4±0.32	11.0
15-17	444	12.1±0.35	11.0

[1]blood lead measured in µg/dL

Importantly, at the time this publication appeared, the Centers for Disease Control criterion for an "elevated blood lead level" (i.e. a "level of concern") was 30 µg/dL, established as such in 1978 (Mahaffey et al. 1982, p. 578).

Race-specific data on blood lead levels in the 2 youngest age groups in the years 1976-1980 revealed the following for white versus black children, adapted from Mahaffey et al. (1982, Table 2, p. 576):

Age Group (years)	# Persons Examined	Mean Blood Lead Level[1]	Median Blood Lead Level[1]
WHITE			
0.5-2	589	15.0±0.56	14.0
3-5	1287	14.9±0.41	14.0
BLACK			
0.5-2	141	20.9±0.96	19.0
3-5	278	20.8±0.55	20.0

[1]blood lead measured in µg/dL

Differences in blood lead levels between white and black children were also present when the data was further stratified by annual family income, as shown in the following table, adapted from Mahaffey et al. 1982, Table 3, p. 577):

Annual Family Income	White # Persons Examined	White Mean Blood Lead Level[1]	Black # Persons Examined	Black Mean Blood Lead Level[1]
< $6,000	256	18.1±0.6	176	22.9±0.9
$6,000-14,999	887	15.3±0.5	163	20.7±0.6

≥ $15,000	690	13.7±0.4	60	17.2±0.8

[1]blood lead measured in µg/dL

Note also the apparent inverse relationship between blood lead levels and annual family income.

Between 1978 and 1990, mitigation efforts reduced or eliminated lead in a variety of products. For example, the amount of lead in gasoline declined by 99.8%, lead solder in food and soft-drink cans was almost eliminated, the use of lead paint was banned, and the use of lead-containing solder in household plumbing was banned. In addition, the country saw an increase in lead paint abatement programs, an increase in lead testing of children, increased public awareness of the issues, and a reduction in the number of occupied dwellings built before 1940 (MMWR, 1994).

Jones et al. (2009) compared blood lead levels (BLL) of children aged 1-5 years and measured in 1988-1991 and in 1991-1994 with blood lead levels measured in 1999-2004. All data originated in NHANES surveys. Between 1988-1991 and 1999-2004 the overall (all races) prevalence of children with blood lead levels ≥ 10 µg/dL decreased from 8.6% to 1.4% an 84% decline. Declines were also observed in the population-based geometric mean BLL of the three primary ethnic groups, non-Hispanic blacks, Mexican-Americans, and non-Hispanic whites, as shown in the following table:

Ethnic Group	Geometric Mean BLL[1] (1988-1991)	Geometric Mean BLL[1] (1999-2004)
Non-Hispanic Black	5.2	2.8
Mexican-American	3.9	1.9
Non-Hispanic White	3.1	1.7

[1]blood lead levels (BLL) measured as µg/dL

Nevertheless, the authors noted that in the NHANES data from 1999-2004 non-Hispanic blacks had the highest percentage of individuals with BLLs ≥ 10 µg/dL, as shown in the following table, adapted from Jones et al. (2009, Table 1, p. e377):

BLL (µg/dL) Geometric Mean	All Races (%)	Non-Hispanic Black (%)	Mexican-Americans (%)	Non-Hispanic White (%)
< 1	14	4	10.9	17.6
1 to < 2.5	55	42.5	61.0	57.1
2.5 to < 5	23.6	36.2	22.1	19.7
5 to < 7.5	4.5	9.4	3.4	3.6
7.5 to < 10	1.5	4.6	1.3	0.8
≥ 10	1.4	3.4	1.2	1.2

Tsoi et al. (2016) report blood lead levels from 1999 through 2014 in the U.S. population using data from the NHANES survey. The authors report the estimated percentages of children (ages 1-5 years) with blood lead levels (BLL) ≥ 5 µg/dL (considered "elevated") during the years 1999-2014 are shown in the following table, adapted from Tsoi et al. (2016, Fig. 2, p. 1217):

Year	% (95% CI)
1999-2000	9.9 (7.5-12.9)
2001-2002	7.4 (5.9-9.4)
2003-2004	5.3 (4.1-6.9)
2005-2006	2.9 (2.1-3.9)
2007-2008	3.1 (2.0-4.8)

2009-2010	2.1 (1.5-3.1)
2011-2012	2.0 (1.0-3.6)
2013-2014	0.5 (0.3-1.0)

The authors conclude that "there has been a continuous decreasing trend in blood lead levels in the U.S. population" (Tsoi et al. 2016, p. 1218).

McClure et al. (2016) examined NHANES data on blood lead levels in children < 6 years of age during the years 2009-2015. They report that "there were significant differences in high blood lead levels based on sex, pre-1950s housing construction quintiles, and PIR<1.25 (PIR is a poverty index), Health and Human Services regions, states, and 3-digit ZIP code areas" (McClure et al. 2016, p. 173).

Summary: Sources of Lead

An individual's blood lead level is a result of exposure to all prior and current sources of lead in the environment over the lifespan of that individual. For children, those sources of lead exposure include but are not limited to dust, soil, water, food, toys, iron status, and paint. The blood lead value measured in any child is also affected by the child's age, mouthing behaviors, socioeconomic situation (including the extent to which lead exists in the plumbing of the child's home), and ethnicity.

Measuring Lead Levels: Capillary versus Blood Measurement Tests

There are two major approaches to measuring lead levels in children: capillary (fingerstick) and venous (blood draw). Both are used extensively in practice and were used to measure lead levels in the plaintiffs in this litigation. The purpose of this section is to compare and contrast the validity and reliability of these two tests. One of the primary concerns about using capillary (fingerstick) tests is that they are more easily contaminated than blood draws (Johnson et al. 1997). As Taylor et al. (2003, p. 217) note, "even low concentrations of contamination can significantly alter the concentration of lead in blood." The source of the contamination is the external environment, e.g. the skin of the individual administering the test as well as the skin of the individual being tested. Several studies of the adequacy of capillary tests for determining blood lead levels have been published (Johnson et al. 1997; Parsons et al. 1997; and Taylor et al. (2003). These studies conclude that, on average, capillary tests overestimate the blood lead levels, largely due to contamination, and they conclude that any capillary test must be confirmed with a venous blood sample. Put another way, capillary blood testing is biased relative to venous blood testing.

For example, Parsons et al. (1997, p. 311) write:

> In conclusion, the fact that capillary BPb is highly correlated with but positively biased relative to venous BPb means that its utilization will induce a 5-10% (7.5-15 μ/L) bias in BPb determinations….provided that the patient's hands have been thoroughly washed beforehand. Lapses in clean technique may lead to an 11-42% bias…"

For example, Taylor et al. (2003, p. 217) write:

> "The findings of this study suggest that until satisfactory skin cleansing and decontamination techniques are identified and evaluated, earlobe sampling should be avoided in the surveillance of occupational blood lead levels."

Johnson et al. (1997, p. 179) write:

> "…all elevated whole blood lead screening results, venous or capillary, should be confirmed with a venous collection before follow-up action is taken."

In sum, capillary blood testing is likely to overestimate blood lead levels. Capillary blood testing is biased but may be useful for initial screening. In any situation, capillary blood tests need to be confirmed with venous tests.

Section 3.2 Cognitive and Behavioral Disorders

Major cognitive and behavioral disorders include intellectual disability, attention-deficit/hyperactivity disorder (ADHD), conduct disorder, and autism. Intellectual disability is a generalized disorder characterized by significantly impaired cognitive functioning. ADHD is a neurodevelopmental disorder characterized by inattention and disorganization, with or without hyperactivity-impulsivity causing impairment of functioning. Conduct disorder is characterized by a pattern of antisocial behaviors. Autism is a neurodevelopmental disorder characterized by severe impairment in reciprocal social interactions (Scott et al. 2015). The prevalence of autism and ADHD in U.S. children is approximately 15% (Lam et al. 2017); cognitive disorders affect approximately 2 to 5% of children (Dooling, 1993).

Section 3.2.1 Background on IQ Measurement in Children

Intelligence tests assess a child's mental abilities and compare them with the abilities of others using a nominal score (Braaten and Norman, 2006). Put another way, a child's IQ score is relative to that of children of the same age. The "others" or the "children of the same age" refer to the children's IQ score in the comparator group, sometimes called the "normative base" (Trahan et al. 2014). An IQ score relates the extent to which a child's mental abilities depart from the average of that comparator group. The average score is 100 on most IQ tests with a standard deviation of 15. The IQ is typically considered to be normally distributed. It follows that 68% of the IQ scores fall with the range of 85 to 115.

<u>IQ Tests for Children</u>

The most-commonly used IQ test is the Wechsler Intelligence Scale for Children. It tests a range of verbal, visual-spatial and problem-solving skills, sometimes referred to as "domains." An overall IQ score is made up of several single factor scores; these scores should be interpreted with caution because there can be variability across domains (Braaten and Norman, 2006).

<u>The Flynn Effect</u>

The "Flynn Effect" is a widely accepted phenomenon documenting a prominent variability in IQ scores over time. Specifically, the Flynn effect refers to the observed rise over time in standardized intelligence tests such that "an individual will likely attain a higher IQ score on an earlier version of a test than on the current version" (Trahan et al. 2014, p. 2). Simply put, the Flynn effect—first discovered by Flynn et al. (1984)—revealed that the IQ scores that comprise the comparator group rise over time, at a rate of approximately 0.31 points per year or 3 points per decade. The scientific community has recommended that corrections for the Flynn effect be implemented at the individual level (Fletcher et al. 2010). A correction involves changing the IQ score to reflect when the comparator group's mean IQ score—the

normative base—was observed. It is that comparator group mean score that changes over time, tending to increase 0.31 points per year. It follows that any study of lead and IQ needs to take the Flynn effect into account. That said, it is not mentioned by any of the plaintiffs' experts in this matter and is ignored in the Lanphear et al. (2005) publication.

Factors other than IQ Affecting Academic Success

There are many factors that play important roles in determining academic success, such as parenting, the quality of education, personal motivation, and exposure to culture and books, as well as IQ (Braaten and Norman, 2006, p. 404-5). A recent study by Breslau et al. (2001) reported that the IQs of urban children—i.e. IQs of children in the City of Detroit, Michigan—regardless of birth weight, declined from age 6 years to age 11 years. The authors suggest that "growing up in a racially segregated and disadvantaged community, more than individual and familial factors, may contribute to a decline in IQ score in the early school years" (Breslau et al. 2001, p. 711).

Individual Variability in IQ Scores

Individual variability in IQ scores is also well appreciated in the scientific community (Ramsden et al. 2011). Whereas there appears to be broad agreement that cognitive skills keep relatively stable over time, it is also true that this stability varies as a function of the age of the sample. Longitudinal (rather than cross-sectional) studies of intelligence over time are uncommon. Those that have been published reveal that IQ tests of young children do not predict later cognitive ability. For example, "in the Fullerton Longitudinal Study correlations between cognitive measures at the age of 17 and preschool measures varied from r = .16 at one year to r = .44 at age 3-5 years" (Schneider et al. 2014, p. 156, citing Gottfried et al. 2006).

> Another example: "Most available studies did not find much longitudinal consistency between intelligence test scores at preschool age and intelligence in later development" (Schneider et al. 2014, p. 156, citing Bishop et al. 2003; and McCall et al. 1972).

Braaten and Norman (2008, p. 405) write that "overall, the general rule of thumb is that the older the child the more stable the IQ. By age 4 years, the correlation with IQ 12 years later is relatively high (r = .77) (citing Neisser et al. 1996). Although many older children show little fluctuation in their IQ scores, research has indicated that a subset of younger children show wide fluctuation in IQ scores. Schneider et al. 2014, p. 156 agree. They note:

> "Most available studies did not find much longitudinal consistency between intelligence test scores at preschool age and intelligence in later development" (citing Gottfried et al. 2009; Bishop et al. 2003; and McCall et al. 1972).

Notably, even older children may show some fluctuations in scores in response to major stressors such as a loss of a parent, divorce, or change in schools. With these possible exceptions, by around age 10 years, IQ scores generally are relatively stable" (Braaten and Norman, 2008, p. 405).

> "Overall, IQ tests are the most reliable and valid instruments used to measure a person's cognitive abilities, but they always should be interpreted within a conceptual framework that does not overstate its implications for the child" (Braaten and Norman, 2008, p. 407).

Fundamental Issues in the Measurement and Interpretation of IQ Scores

There are some fundamental issues in the measurement and interpretation of IQ scores and changes in IQ scores that require attention.

For example, a scientifically valid assessment of the change in IQ potentially due to an exposure requires that an individual be tested before the exposure and after the exposure. As Morris et al. (1999, p. 789) write in their critical discussion of methodological issues in the study of cognitive decline:

> Measurement of cognitive decline "requires measurement of change in function over time rather than a single static measurement."

Most (if not nearly all) studies of lead exposure and IQ scores (and other cognitive and behavioral outcomes) rely solely on single measures usually years after the exposure or at the same time as the exposure. These studies do not provide information—actual measures—on IQ before lead exposure and IQ after lead exposure. The methodological problem is clear. It is scientifically inappropriate to assign to causation to lead exposure for a change in IQ that was never measured as such.

Another example involves the fundamental nature of IQ scores. These scores are a different type of measurement than measures like inches or quarts or ounces. Inches, quarts, and ounces are ratio scales where zero means zero and 100 units are twice 50 units. Intelligence test scores, however, use interval scales and have meaning only relative to the scores of other people of the same age and sex. As Haier (2014, p. 1) writes:

> "…someone with an IQ score of 130 is not 30% smarter than someone with an IQ score of 100. A score of 130 puts someone in the highest 2% of the population whereas a score of 100 is at the 50th percentile. A change from an IQ score of 100 to 103 is not the same as a change from 133 to 136. This makes interpretation of intelligence test score changes impossible."

Note the word "impossible." In other words, the notion that simple regression techniques like those used in the study of lead exposure and IQ by Lanphear et al. (2005), besides not controlling for the Flynn Effect and not controlling for many known and suspected neurotoxins, have a fundamental flaw namely that attributing changes to the hypothesized effects of lead, even when statistically significant, is seriously problematic. As Haier (2014, p. 2) writes:

> "When change scores are used, it is important to identify individual differences even within a group where the average change score statistically increases after an intervention."

Not everyone in the "intervention" (i.e. exposed) group will have had a change in intelligence, independent of the IQ score they record. In essence, Haier (2014) concludes that even small statistically significant changes in test scores are not sufficient proof that general intelligence has changed. Add to this fundamental uncertainty the notion that a 5-point IQ change in an individual is considered clinically insignificant (Breslau et al. 2001, p. 716).

Finally, it is important to remember that a child's performance on an IQ test can often be affected by factors other than the exposure under study, including educational attainment, cultural experiences, language use, prior testing experience, emotional and physical states, the testing environment, and measurement error (Morris et al. 1999).

Summary: Implications for the Study of IQ and Lead Exposure in Children: Flint, Michigan

Important implications of the foregoing description and discussion for examining the validity and reliability of studies of IQ and lead exposure include the following:

1. Small changes in IQ (e.g. 5 points) are considered clinically insignificant.

2. Failure to measure the change in IQ for each individual study participant makes any claim regarding the relationship of lead exposure to change in IQ scientifically challenging if not invalid. It follows that failure to measure the change in IQ (or other measures of intellectual achievement) pre and post lead exposure in the Flint, Michigan children makes any claim regarding individual causation challenging if not invalid.

3. Failure to adjust for the Flynn effect in studies of lead exposure and cognitive change is a source of bias in those studies.

4. Even if it were assumed that IQ changes (specifically, decrements) are causally associated with lead exposure in groups (populations) of children, it is scientifically inappropriate to claim that each and every child in the group will have had a change in IQ attributable to lead exposure.

Section 3.2.2 Inattention, Hyperactivity, and Impulsivity Often Associated with Attention-Deficit/Hyperactivity Disorder

Attention-deficit/hyperactivity disorder (ADHD) is one of the most common childhood disorders, with estimated prevalence of 3-8% (Kim et al. 2013; Nigg et al. 2008; Braun et al. 2006). ADHD is defined as a disorder marked by an ongoing pattern of behaviors involving inattention, hyperactivity, and/or impulsivity that interfere with normal functioning or development (NIMH website).

> By "inattention" it is meant that a child lacks persistence, has difficulty maintaining focus, is disorganized and typically wanders off assigned tasks. Importantly, these features are not due to the child being defiant or lacking comprehension.
>
> By "hyperactivity" it is meant that the child moves constantly including those situations in which movement is not appropriate. Excessive talking, tapping, or fidgeting are common features.
>
> By "impulsivity" it is meant that the child makes actions too hastily without thinking about them. These actions may be taken with a high harm potential or accompanying a desire for immediate reward.

There are three recognized subtypes of ADHD: (1) predominantly hyperactive (ADHD-PH), (2) predominantly inattentive (ADHD-PI), and (3) combined (ADHD-C). Most children exhibit symptoms of ADHD-C, the combined subtype (Nigg et al. 2008; NIMH website).

Importantly, the behavioral symptoms commonly found in ADHD are often evaluated even when a diagnosis of ADHD has not been made. These related behavioral symptoms—inattention, hyperactivity, and impulsivity—feature prominently in the discussion and evaluation of evidence regarding exposure to lead and behavioral outcomes in children.

Risk factors for Inattention, Hyperactivity, and Impulsivity in ADHD

Many factors are known to increase the risk of inattention, hyperactivity, and impulsivity, the behavioral symptomatology of ADHD. Genetic factors play the most important etiologic role; heritability accounts for 60-80% of ADHD and its symptoms (Azeredo et al. 2018; Van Dongen et al. 2019; Kim et al. 2013; Choi et al. 2016; NIMH website). Other factors than genes can also increase risk. These will be discussed more fully later in this report in terms of their confounding effects on the relationships between exposure to lead and behavioral outcomes consistent with ADHD. They are listed here:

- Male gender (Ji et al. 2018)
- Preterm birth (Allotey et al. 2018; Franz et al. 2018; Nielsen et al. 2019)
- Maternal obesity (Cortese and Tessari, 2017)
- Diet/Nutrition (Del Ponte et al. 2019)
- Maternal smoking (Dong et al. 2018; Huang et al. 2018)
- Air Pollution including particulate matter and NO_2 (Donzelli et al. 2019; Fuertes et al. 2016; Sentis et al. 2017; Thygesen et al. 2020)
- Maternal thyroid dysfunction (Drover et al. 2019; Fetene et al. 2017)
- Acetaminophen use during pregnancy (Gou et al. 2019)
- Vitamin D deficiency (Khoshbakht et al. 2018)
- PBDE exposure (Lan 2017)
- Hypertension of pregnancy (Maher et al. 2017)
- Phthalate exposure (Shoaff et al. 2020)
- Exposure to PCBs (Verner et al. 2015)
- Maternal alcohol consumption (Wetherill et al. 2018)
- Cesarean delivery (Zhang et al. 2019)
- High levels of exposure to other environmental pollutants such as Hg, Cd, Mn, Pb (Kim et al. 2013; Chan et al. 2015; NIMH website)

Note that this list was created from the National Institute of Mental Health (NIMH) website and from English-language systematic reviews and meta-analyses identified in a PubMed search undertaken on October 10, 2020 using search terms “ADHD” and “children” and “etiology” and “epidemiology.” Filters included having been published in the past 5 years. There were 64 publications identified before assessing relevance.

Diagnosis of ADHD and Related Behavioral Symptomatology

> “Diagnosis of ADHD requires a comprehensive evaluation by a licensed clinician, such as a pediatrician, psychologist, or psychiatrist with expertise in ADHD. For a person to

> receive a diagnosis of ADHD, the symptoms of inattention and/or hyperactivity-impulsivity must be chronic or long-lasting, impair the person's functioning, and cause the person to fall behind typical development for his or her age. The doctor will also ensure that any ADHD symptoms are not due to another medical or psychiatric condition. Most children with ADHD receive a diagnosis during the elementary school years. For an adolescent or adult to receive a diagnosis of ADHD, the symptoms need to have been present before age 12" (NIMH website).

A wide variety of tests for symptoms of ADHD, e.g. inattention and hyperactivity, are available. One of the most common types of tests is to ask parents and teachers if these symptoms have been observed in the child under examination; typically, these are questionnaires with rating scales often based on the Statistical Manual of Mental Disorders published by the American Psychiatric Association. It is important that both parents and teachers provide independent assessments (Pelham et al. 2005).

Summary: Implications for Studying Behavioral Disorders and Lead Exposure: Flint, Michigan

In the absence of independent assessments by both parents and teachers of a child's behavioral problems before and after the Flint Water Switch, attributing lead exposure as the cause is fraught with uncertainty. Similarly, in the absence of objective tests, the same conclusion applies. Finally, the failure of investigators to fully adjust for the known risk factors for behavioral disorders (broadly defined) negatively affects the validity and reliability of epidemiological studies examining the hypothetical relationship between lead exposure and these outcomes. I turn now to what is known and not known about lead exposure in Flint, Michigan.

PART FOUR: LEAD and FLINT, MICHIGAN

Section 4.1 Studies of Flint Michigan Children's Blood Lead Levels

The purpose of this section is to describe blood lead levels in children before, during, and after the so-called "Flint Water Switch," i.e. the increase in lead content of drinking water in Flint, Michigan that occurred from April 25, 2014 through October 15, 2015 (Kennedy et al. 2016). According to the Centers for Disease Control investigators, "residents of Flint, Michigan were affected by changes in drinking water quality after their water source was switched from the Detroit Water Authority (DWA), sourced from Lake Huron, to the Flint Water System (FWS), sourced from the Flint River" (Kennedy et al. 2016, p. 650). These same investigators noted that a water advisory was issued to the residents of Flint, Michigan on January 2, 2015 (Kennedy et al. 2016).

Several studies of blood lead levels (BLLs) in Flint, Michigan children have been published and are briefly described below, in chronological order.

Hanna-Attisha et al. (2016) is an early retrospective ecologic study of blood lead levels in children living in Flint, Michigan before and after the Flint Water Switch in 2014-5). These investigators reported on the percentage of children with elevated blood lead levels (where "elevated" means > 5 µg/dL) stratified by neighborhoods in Flint, Michigan. The authors reported that the incidence of elevated blood lead levels increased from 2.4% to 4.9% after the water switch; some neighborhoods experienced an increase to 6.6%. The authors provide a discussion of limitations of their study but fail to point out the fact that their study is subject to the ecologic fallacy, is best considered a hypothesis-generating study (rather than a hypothesis-testing study). For more details on the "ecologic fallacy, see Appendix D of this report. Briefly, a study designed in this manner cannot determine which child living in Flint, Michigan was exposed to an increased lead level in their drinking water. As such, causal claims from this study, by design, are invalid. The authors also failed to control for seasonal effects of lead exposure, the fact that children may not have been exposed if they began drinking bottled water, and the advisory described in more detail in the study immediately below (i.e. Kennedy et al. 2016). In the end, the study provides, at best, a rationale for further study with a better design and better methods.

Kennedy et al. (2016) reported an analysis of BLLs in Flint, Michigan children (aged < 6 years) during four time periods: before the change from DWA to FWS (April 25, 2013 – April 24, 2014), after the change and before the water advisory (April 25, 2014 – January 2, 2015), after the advisory and before the switch back to DWA (January 3, 2015 – October 15, 2015) and after the switch back to DWA (October 16, 2015 – March 16, 2016). BLLs were obtained from the Michigan blood lead surveillance system that targets children living at or below the poverty level as well as children enrolled in Medicaid. Venous blood tests were preferred over capillary blood tests. The authors noted that "if a BLL of ≥ 5 µg/dL was reported during a given period, no subsequent blood lead tests from that child were included in the analyses" (Kennedy et al. 2016, p. 650).

Numbers and percentages of elevated (≥ 5 µg/dL) blood lead test results in children (< 6 years) by the four assessment periods are shown in the table below, adapted from Kennedy et al. (2016, Table 1, p. 652):

BLL Levels	Before Switch from DWA to FWS No. (%)	After Switch to FWS Before Advisory No. (%)	During Water Advisory No. (%)	After Switch from FWS back to DWA No. (%)
≥ 5 µg/dL	74 (3.1)	84 (5.0)	78 (3.9)	48 (1.4)
5-9 µg/dL	59 (2.5)	71 (4.2)	68 (3.4)	37 (1.1)
10-14 µg/dL	9 (0.4)	10 (0.6)	6 (0.3)	4 (0.1)
15-19 µg/dL	2 (0.1)	2 (0.1)	0 (0)	4 (0.1)
20-39 µg/dL	4 (0.2)	1 (0.1)	4 (0.2)	2 (0.1)
≥ 40 µg/dL	0	0	0	1 (< 0.1)

The data was stratified by various factors, including but not limited to ethnicity (African American, White, Other/Unknown) and age (< 1, 1-2, 3-5) as shown in the following table, adapted from Kennedy et al. (2016, Table 1, p. 652):

Factor	Before Switch from DWA to FWS No. (%) ≥ 5 µg/dL	After Switch to FWS Before Advisory No. (%) ≥ 5 µg/dL	During Water Advisory No. (%) ≥ 5 µg/dL	After Switch from FWS back to DWA No. (%) ≥ 5 µg/dL
Age Group (yrs)				
< 1	1 (0.8)	5 (3.4)	3 (2.0)	3 (0.8)
1-2	58 (3.7)	59 (5.7)	57 (4.6)	26 (2.0)
3-5	15 (2.1)	20 (3.9)	18 (3.0)	19 (1.2)
Race				
African American	38 (2.8)	56 (5.5)	47 (4.1)	29 (1.4)
White	25 (4.4)	24 (5.5)	23 (4.0)	16 (1.8)
Other/Unknown	11 (2.2)	4 (1.7)	8 (3.1)	3 (0.7)

The authors also reported multivariable adjusted odds ratios (aORs) with 95% confidence intervals (95% CI) comparing odds of elevated blood lead levels (≥ 5 µg/dL) among children aged < 6 years by selected covariates as shown in the following table, adapted from Kennedy et al. (2016, Table 3, p. 653):

Period	AOR[1] (95% CI)
Before Switch to FWS from DWA	Referent
After Switch to FWS before Advisory	1.46 (1.06-2.10)
Advisory	1.28 (0.92-1.76)
Switch from FWS to DWA	0.75 (0.51-1.12)

[1]adjusted for age and season

The authors concluded that the "adjusted probability of having BLLs ≥ 5 µg/dL was 46% higher during the period after the switch from DWA to FWS (and before the January 2, 2015 water advisory) than during the period before the water switch to FWS" (Kennedy et al. 2016, p. 650).

The authors note that they did not control for sources of lead other than that in the drinking water. Indeed, they note that "there might be multiple sources of early childhood lead exposure in areas with houses built before lead paint use in the United States was banned in 1978" (Kennedy et al. 2016, p. 653). Houses in Flint, Michigan tend to be older and, as a result, have lead-based paint. In fact, the city of Flint has recently implemented a program entitled "Lead-Based Paint Hazard Control" through which "Residents can have lead-based paint hazards remediated from their homes at NO COST through a special grant program being operated by the City of Flint and funded through the U.S. Department of Housing and Urban Development" (see www.cityofflint.com/lbphc).

Laidlaw et al. (2016) is the report of a study of the seasonal patterns of Flint, Michigan children's blood lead levels (BLLs) with a focus on the extent to which soil lead contributed to observed increases during the years 2010-2015. The authors conclude that the seasonality of BLLs observed in these children during the time when the Flint water source was changed was due at least in part to the well-established seasonality of lead exposure from lead-contaminated urban soil. They write that the "blood lead peaks in the third quarters of 2014 and 215 may have been driven by a combination of lead from resuspended soil-derived dust and lead newly released from water lines" (Laidlaw et al. 2016, p. 9). Their conclusions are based in part upon the fact that "soil lead and its seasonal suspension and deposition into homes is a major contributor to chronic lead exposure in the United States" (Laidlaw et al. 2016, p. 9).

Seasonality of blood lead levels in children is well known. Using data from the Centers for Disease Control and Prevention, Laidlaw et al. (2016) show evidence of the phenomenon for the United States with a clear increase in the numbers of cases of elevated blood lead levels ≥ 10 µg/dL beginning in May (2012) through August (2012) and then a clear decrease through December (2012). Several other studies from the U.S. are cited confirming this phenomenon including the study by Zahran et al. (2013). See also Zahran et al. (2017, p. 170) who note that throughout their study periods (i.e. 2013-2016) they observed the "same seasonal pattern to BLLs in Flint (and in Genesee County) children."

As Laidlaw et al. (2016) discuss, the intense media attention given to the Flint, Michigan water switch and especially their attention on the observation that Flint children's blood lead levels increased (from 2.4% to 4.9%) between 2013 and 2015 (citing Hanna-Attisha et al. 2016) failed to fully account for the well-established seasonality patterns of summertime exposure to lead-contaminated dust from soil. As Laidlaw et al.'s Figure 2 (2016, p. 6) and Figure 3 (2016, p. 7) clearly show—see below—there have been consistent peaks in the percent blood lead level ≥ 5 µg/dL in Flint children during the summer months (i.e. the third quarter or Q3). Importantly, these same figures show that the peak for Flint children in 2014 Q3 lessened in 2015 Q3. The same pattern can be seen, indeed, from 2010 through 2015 for Genesee County (including Flint) and the state of Michigan in the Laidlaw et al. (2016 Fig. 3, p. 7) paper.

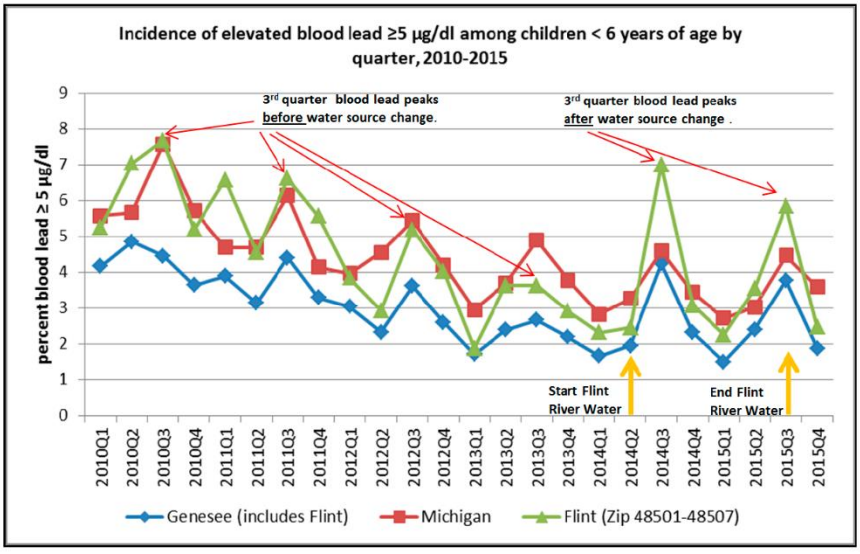


Figure 2. Incidence of blood lead ≥5 µg/dL among children <6 years of age by quarter from 2010 to 2015 (source of data—MDEQ, 2016 [59]).

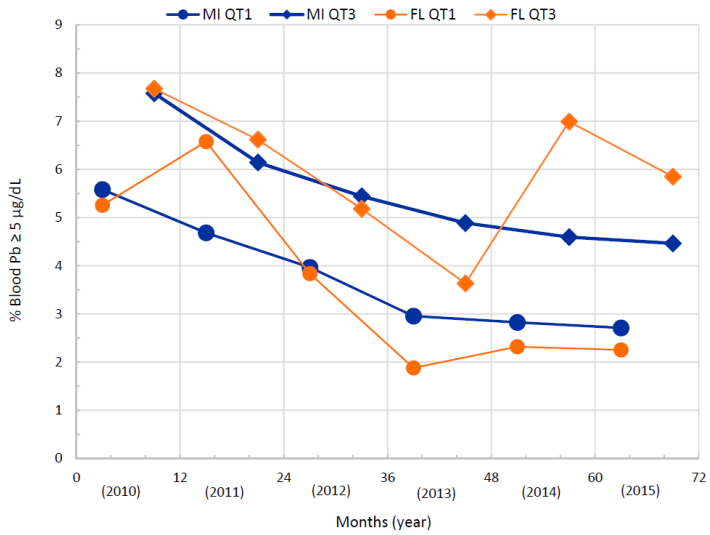


Figure 3. Seasonal differences between Quarter 1 and Quarter 3 for six years of percent blood lead ⩾5 µg/dL for the children of the State of Michigan and Flint area children. The data in this figure is the same as shown in Figure 2, but it emphasizes the seasonal differences in the blood lead.

In summary, soil contaminated with lead plays an important role in the seasonal variation of blood lead levels in children, including those living in Flint, Michigan during the years 2014-2015.

Gomez et al. (2018) evaluated blood lead levels (BLLs) in Flint, Michigan children during the years 2006-2016. The authors identified 15,817 blood lead levels of children ≤ 5 years of age, with a mean age of 2.4 years. Beginning in 2013, most (88-94%) of the BLLs were measured in children insured through Medicaid. The authors calculated geometric mean blood lead levels (GM BLLs) and calculated the change in BLLs ≥ 5 µg/dL using pairwise comparisons by year. GMs are often used in these circumstances to provide a better estimate of central tendency when the data has a "long tail" at the upper end of the distribution. Geocoded addresses of the individual study participants were used to ensure that those participants resided in Flint, Michigan. The authors controlled for seasonality and used a Bonferroni adjustment to account for the multiple year comparisons. Results showed that from 2006 to 2016 there was a 72.9% decrease in the percentage of BLLs ≥ 5 µg/dL with a steady decline in the percentage of children above the reference value. In 2006, that percentage was 11.8% while in 2016 the percentage of children above the reference value was 3.2% (95% CI: 6.7-10.4, $p < 0.001$). Although this same percentage increased during the years 2014 (3.3%) and 2015 (3.7%), these differences were not statistically significant when the Bonferroni adjustment was applied. See Gomez et al. (2018, Fig. 1, p. 160).

When GM BLLs were analyzed, a similar pattern was observed; see Gomez et al. (2018, Fig. 2, p. 161). There has been a steady decrease in GM BLLs during the years 2006 through 2016 with an increase between 2010 and 2011 (statistically significant) as well as an increase from 1.19 to 1.30 µg/dL between 2014 and 2015, representing an increase of 0.11 µg/dL during the Flint River water years of 2014-2015.

The authors conclude that their "findings suggest that public health efforts to reduce BLLs of young children in Flint have been effective over the 11-year period studied" (Gomez et al. 2018, p. 163).

In an editorial accompanying the Gomez et al. (2018) publication, **Banner (2018)** noted that the authors provided the reader with a "straightforward presentation of the data" in the absence of what he called

"toxicohistrionics," to describe what he perceived to be public and professional overreactions to the Flint Water situation.

The implications of the results of Gomez et al. (2018) are important for the Flint children. Consider the following problem: if the average (i.e. geometric mean) increase in blood lead levels (BLLs) among Flint children was approximately 0.11 µg/dL during the Flint Water Switch (FWS), then what change in IQ across that same population would be predicted (i.e. estimated) assuming that exposure to lead causes a negative change in IQ? An answer to this question can be found using the reported results from the pooled study of Lanphear et al. (2005) described earlier in this report. In that analysis, an increase in blood lead level from 2.4 to 10 µg/dL was associated with a 6.9 point IQ decrement. Put another way, if a 7.6 µg/dL (i.e. 10 – 2.4) is associated with a 6.9 IQ decrement, then a 0.11 µg/dL increase would be associated with a 0.1 point decrement in IQ. This calculation assumes an approximate linear relationship.

As Braaten and Norman (2006, p. 405) write: "research has indicated that a subset of younger children show wide fluctuations in IQ scores." Even common sense says that a change of 0.1 is not a wide fluctuation. By age 10, "IQ scores generally are relatively stable." However, by relatively stable, the test results are within a few points not within a tenth of a point. Indeed, IQ scores are always reported as whole numbers and not with decimals. In any case, a 0.1 decrement in an IQ score is clinically insignificant and uninterpretable.

Gomez et al. (2019) analyzed the BLLs from Flint, Michigan children aged ≤ 5 years using three time periods: Period I (April 25, 2006 to October 15, 2007), Period II (April 25, 2012 to October 15, 2013), and Period III (April 25, 2014 to October 15, 2015). These three periods were chosen to represent the earliest time period before the Flint water switch (FWS), the time period closest to the FWS, and the time period coinciding with the FWS. Results revealed that the percentage of Flint children with BLLs ≥ 5 µg/dL were no different during the FWS than during the period immediately preceding the FWS, as shown in the following table, adapted from Gomez et al. (2019, Table 2, p. 793):

	Period I % (N)	Period II % (N)	Period III % (N)	Period I v II p-value 95% CI	Period I v III p-value 95% CI	Period II v III p-value 95% CI
% ≥ 5 µg/dL	10.6 (2095)	3.3 (1834)	3.9 (1734)	P<0.001 (5.7-8.8)	p=.002 (5.0-8.2)	p=0.30 (-1.9-0.57)

Summary

Measured blood lead levels in young children (≤ 5 years of age) residing in Flint, Michigan during the switch from drinking water provided by the Detroit Water Authority (DWA) to the Flint River (FRW) revealed weak temporary elevations in average BLLs consistent with random variation and with the well-recognized seasonality of blood lead levels due to lead-contaminated soil. The average increase in BLLs among the Flint children during the (FWS) was approximately 0.11 µg/dL, an increase that could be associated with changes in neurodevelopmental outcomes (e.g. IQ) so small as to be clinically insignificant and uninterpretable.

Section 4.2 Data on Class Certification Plaintiffs

Blood lead levels for the representative class certification plaintiffs were provided in various documents, including medical records and addendums to legal documents, including interrogatories.[3] These are shown in the table below.

Plaintiff (Date of Birth)	Date	Lead Level µg/dL
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

Summary of Blood Lead Levels of Class Certification Plaintiffs

The data on blood lead levels of the class certification plaintiffs reveals that no level is above 5 µg/dL at any time before, during, or after the Flint Water Switch. Given that a causal association between blood lead and exposure and IQ (as one example of an outcome) has not been established at blood lead levels < 5 µg/dL, these plaintiffs have not experienced clinically significant decreases in IQ as a result of the Flint Water Switch. Note also that in the source documents for these blood lead levels, there is little or no information on the other 20 factors that could have affected blood lead levels in these plaintiffs. As importantly, with one possible exception ([redacted]), the blood lead levels reveal no increases in blood lead levels consistent with having been exposed to increased lead during the Flint Water Switch. Instead, the remainder of the blood lead levels reveal constant levels (e.g. [redacted]), single values ([redacted]), decreases ([redacted]) or what appears to be random variability ([redacted]). In fact, any differences observed between successive blood lead levels for any given individual in this table could be explained by random variability, measurement error, or other factors—i.e. other than lead—known to influence blood lead levels.

[3] For details regarding the source documents used to create this table, see Section 8.3 "Additional Documents."

PART FIVE: GENERAL CAUSATION

In Section 5.1, I briefly describe the methods used to assess evidence for the purpose of making claims of general causation. A more complete description of these methods can be found in Appendix D. In Section 5.2, I apply these methods to the evidence on the nature of the relationship between exposure to lead in children and cognitive and behavioral outcomes. In Section 5.3, I examine the issue of a threshold for the effects of lead on cognitive and behavioral outcomes. In Section 5.4, I apply these analyses to the population of Flint, Michigan.

Section 5.1 Methods of General Causation: A Brief Overview

In current practice, claims about disease causation emerge from the application of causal inference methods to bodies of scientific evidence. These methods include: (1) the general scientific method, (2) epidemiological methods, commonly described in terms of study design types such as cohort studies, case-control studies, and ecologic studies, to name some examples, and (3) methods of research synthesis. Of the methods of research synthesis, the systematic review—described earlier in this report—is the appropriate and scientifically reliable and valid method to employ. Within this method, a research question is posed, published studies designed to answer this question are identified through systematic literature searches, and these same studies are described. Interpretation of the body of evidence so collected can then proceed.

In this legal matter, the following [4]questions require a systematic approach:

(1) does exposure to lead in childhood adversely affect cognitive abilities, and if so, to what extent—if any—were the children of Flint, Michigan who were exposed to lead in drinking water during the so-called Flint Water Switch adversely affected in terms of cognition?

(2) does exposure to lead in childhood adversely affect behavior, and if so, to what extent—if any—were the children of Flint, Michigan who were exposed to lead in drinking water during the so-called Flint Water Switch adversely affected in terms of behavior?

Section 5.2 Assessment of General Causation: Lead and Cognitive/Behavioral Outcomes

A detailed description of the systematic literature search for studies examining the relationship between exposure to lead in childhood and cognitive outcomes can be found in Appendix C. Similarly, a detailed description of the systematic literature search for studies examining the relationship between exposure to lead in childhood and behavioral outcomes can be found in Appendix C. The studies identified by these searches are also described in detail in Appendix C.

Briefly, systematic searches of two databases (PubMed and Web of Science) were undertaken using search terms “lead” and “exposure” and “health effects” including but not limited to neurodevelopmental outcomes (cognitive and behavioral). The searches were designed to identify

[4] Note that there are additional general causation issues described, discussed, and interpreted in Appendices, G, H, I, and J. These involve lead and hypertension in children as well as lead and renal disease (and renal dysfunction) in children as well as the occurrence of cardiovascular disease and essential tremor in adults related to childhood lead exposure.

epidemiological studies designed as cohort, case-control, and cross-sectional. Appendix C provides descriptions of the cohort studies of lead exposure in children and neurodevelopmental outcomes identified by these searches.

It is important to point out that exposure to lead resulting in long-term blood lead levels greater than 10 µg/dL has been causally linked to adverse cognitive and behavioral outcomes. Whether these same adverse outcomes occur and can be attributed to lower blood lead levels (i.e. < 10 µg/dL) has been a subject of debate and interest in the scientific and public health communities. The relatively large number of epidemiological studies examining these hypotheses—as described in Appendix C—speaks to the need for more and better studies on this important topic.

Of particular concern is the lack of attention to the well-established alternative causes of adverse cognitive and behavioral outcomes in the published epidemiological studies of low-dose lead. By "lack of attention" I mean the failure on the part of investigators to effectively control for these many alternative causes (and risk factors) through restriction and/or statistical adjustment. I turn to a systematic assessment of this issue in the next section of this report.

Lack of Adjustment for Causes and Risk Factors of Adverse Neurodevelopmental Outcomes in the Epidemiologic Studies of Low-Level Lead

A careful investigation of the efforts by the authors of studies described in Appendix C to control (i.e. adjust) for the many known causes and risk factors of neurodevelopmental outcomes reveals that only a small fraction of the studies of lead and neurodevelopmental outcomes adjust for these factors. The following table shows how many studies—out of the 50 studies (42 cohort studies of lead and intelligence and 8 cohort studies of behavior and intelligence)—adjusted for each of the 17 factors at issue:

Confounding Factor	# of Studies (%)	Comments
Methylmercury or Hg	2 (4%)	Plusquellec et (2010) Inuits Canada Kim et (2016)
Arsenic	3 (6%)	Rodrigues et (2016) Bangladesh Parajuli et (2013) Nepal McDermott et (2011) S. Carolina
PCBs	1 (2%)	Ethier et (2015) Arctic Quebec
Toluene	0 (0%)	
Manganese	2 (4%)	Rodrigues et (2016) Bangladesh Henn et (2012) Mexico
Fluoride	0 (0%)	
Chlorpyrifos	0 (0%)	
Pesticides	1 (2%)	Huang et (2012)
Dichlorophenyltrichloroethane	0 (0%)	
Tetrachlorethylene	0 (0%)	
PBDE	0 (0%)	
Air Pollution	0 (0%)	
Maternal Smoking	14 (28%)	Rodrigez et (2016); Liu et (2013); Jedrychowski et (2008, 2009); Canfield et (2003, 2004); Jusko et (2018); Baghurst et (1994); Huang et (2012); Mazumdar et (2011); Chandramovli et (2009); Taylor et (2015); Kim et (2016); Bellinger et (2014); Taylor et (2017)

Maternal Alcohol	6 (12%)	Plusquellec et (2010); Braun et (2012); Huang et (2012); Mazumdar et (2012); Plusquellec et (2017); Ethier et (2015); Taylor et (2017)
Maternal Obesity	4 (8%)	Vigeh et (2014) Parajuli et (2013) Bellinger et (2014) Choi et al. (2016)
Prenatal Care	1 (2%)	Bellinger et (1994)
Preterm Birth	5 (10%)	Claus et (2012); Evens et (2015) Blackowicz et (2016); Choi et (2016); Kim et (2016)

The evidence is clear. Adjusting for these confounders has rarely been attempted at the individual factor level much less attempting to adjust for all known confounders. Seven (7) of the seventeen (17) factors have not been adjusted for in any of the 50+ studies mentioned above and described in more detail in Appendix C. For the remaining ten (10) factors, only 2—maternal smoking and maternal alcohol—have been adjusted for in more than 10% (>5) studies.

It is scientifically appropriate to conclude that the observed relationship between exposure to lead (measured with blood lead levels) and neurodevelopmental outcomes cannot be considered established in the absence of effective adjustment for well-established causes and risk factors—confounders—of the association between lead exposure and neurodevelopmental outcomes. This conclusion is particularly important for studies of blood lead levels ≤ 5 µg/dL where data on the relationship between lead and neurodevelopmental outcomes is sparse and measurement error is an important concern as discussed by Caldwell et al. (2017). However, precisely the same issue—failure to control for confounding—also affects studies examining the potential relationship of lead exposure at levels greater than 5 µg/dL, i.e. those studies examining children with blood lead levels between 5 µg/dL and 10 µg/dL. The methodological maxim is clear here: as an association gets weaker, the impact of uncontrolled confounding becomes more important. Indeed, this is the main point of the Bradford-Hill criterion of "strength of association" as discussed in more detail later in this report.

In addition, it is important to point out that some of the studies described above were those included in the Lanphear et al. (2005) analysis described in more detail earlier in this report.

Section 5.3 The Issue of a Threshold for the Effect of Lead on Neurodevelopmental Outcomes

A scientific issue that requires careful consideration is that of a "threshold" for the hypothesized effects of exposure to low levels of lead (i.e. below 10 µg/dL). In the scientific community it is commonly held that a threshold has not been identified. However, the fact that a threshold has not been identified does not mean that a threshold does not exist or, alternatively, that the effects of low levels of lead are clinically relevant. Representative examples of the scientific community's beliefs on the topic of a threshold for the effects of lead on neurodevelopmental outcomes are shown below.

Lanphear et al. (2005, p. 898) report the following:

> "In this pooled analysis, we found evidence of lead-related intellectual deficits among children who had maximal blood lead levels < 7.5 µg/dL. Indeed, we found no evidence of a threshold."

Flannery et al. (2020, p. 2) write: "Research has not identified a threshold level below which neurodevelopmental effects are not expected to occur" citing (CDC, 2018; EPA, 2013; EFSA, 2010; FDA, 2018b; JECFA, 2011).

ATSDR (2019, p. 3) citing CDC (2018b) states that "no safe blood lead level in children has been identified." Similarly, this report notes (ATSDR, 2019, p. 5): "No threshold for these effects has been identified (i.e., no safe level has been identified)."

Note that "no evidence of a threshold" and "no safe level" are used as interchangeable in the ATSDR 2019 report. However, these two concepts are not identical. Safety is a regulatory consideration and not a scientific issue. Typically, "safety" levels are set much below the level of least effect.

Of importance here is the extent to which the scientific studies demonstrate a threshold (or not). The Lanphear et al. (2005) study and the re-analysis by Crump et al. (2013) are informative. The scatter plot of data used by Lanphear et al. (2005) is shown in Crump et al. (2013, Fig. 2, p. 794). There is a paucity of data below 5 µg/dL and a key conclusion from Crump (2013, p. 798) who state:

> "…the question of whether or not a threshold of exposure exists can never be answered definitively by a statistical analysis. There will always be statistical uncertainly that will make it impossible to rule out a nonzero response (either positive or negative) at any exposure, particularly at low exposures."

Crump et al. (2013, p. 798) also conclude:

> "…it is possible to use a statistical analysis to set an upper bound for any threshold that may exist. The present analysis suggests an association of BPb with IQ at concurrent BPb levels as low as 5 µg/dL. Given that, if a threshold for the effect of BPb with IQ does exist, our analysis suggests that that threshold would be below a concurrent BPb of 5 µg/dL or below a peak BPb of 7 µg/dL."

Simply put, the fact that no threshold has been identified does not mean there is no threshold. The precise location of a threshold for the effects of low-level lead exposure is unknown and is as likely to be somewhere between 5 µg/dL and 7 µg/dL (peak) as it is to be somewhere below 5 µg/dL. There is, in other words, high uncertainty about its location. However, it is known that the evidence for neurodevelopmental effects in children below 5 µg/dL is based primarily on extrapolation (modeling with sparse or nonexistent data) and failure on the part of investigators to control for factors known to be associated with or to cause neurodevelopmental disorders.

More on Thresholds and the Plaintiffs' Experts in this Litigation

The idea of a threshold has been mentioned extensively by many of the plaintiffs in this litigation and by several organizations in official reports as well as by authors of published scientific papers as noted earlier in this report. Clarification of the term "threshold" is, therefore, important.

The term "threshold" can be used in two different ways: (1) in mechanistic terms to mean the least amount of a compound—in terms of the number of molecules of that compound—that has an effect at the molecular or cellular level and (2) to mean the least amount of observed exposure that is associated

with an observed effect at the population level sometimes called the "no effect" level. Put another way, it is important to distinguish between threshold at two different but interconnected levels of organization: (1) the molecular level and (2) the population level.

Keeping these two meanings distinct is important because the plaintiffs and, to some extent, the published papers and reports confuse the two or, more precisely, conflate the two, using them interchangeably. In most cases, it appears that the authors are focusing on the idea that there is no least amount of exposure at the population level that can cause a neurodevelopmental effect. When they say "no threshold" they mean "any level of exposure" however small, causes a neurodevelopmental effect in people (e.g. children). Whether this version of a "threshold" equates to a single molecule of the exposure having an effect at the molecular level is unclear. Indeed, the authors of many publications on lead and intelligence argue that the effect of lead increases when the change in lead levels are at their lowest levels despite the fact that evidence at those very low levels is scanty or nonexistent.

Section 5.4 Some Special Methodological Considerations in Assessing the Reliability and Validity of Studies of Lead Exposure and Neurodevelopmental Outcomes in Children

Study Design

If the hypothesis of interest—as it is in this matter—is that a change in blood lead level increases the risk of a change in IQ (or any other neurodevelopmental outcome), then what study design best tests this hypothesis?

1. Measure IQ and blood lead level (BLL) at a single point in time for individuals that are participants in an epidemiological study and then analyze the data using regression revealing how the measured IQs in the study population are distributed relative to the measured blood lead levels.

2. Measure BLLs at time t_0 and then wait several years and measure IQs at time t_1 and then analyze the data using regression revealing how the measured IQs in the study population are distributed relative to the measured BLLs.

3. For a population of children participating in an epidemiological study, measure BLLs and IQs at time t_0 and then again at time t_1 and then analyze how the change in BLLs ($t_1 - t_0$) predicts the change in IQs ($t_1 - t_0$). For example, one could compare how IQ changes in those individuals whose BLLs decreased relative to those individuals whose BLLs stayed relatively constant versus those individuals whose BLLs increased over the time period $t_0 - t_1$. This is an example of a longitudinal study with repeated measurements of exposure and outcome.

4. Measure BLLs and IQs at time t_0 and then again at time t_1 for a group known to have been exposed to an excess of lead (however that occurred) during the time period t_0 to t_1 and compare this data to measures of BLLs and IQs at both time periods for a group not so exposed. This is an example of a longitudinal study with repeated measurements of exposure and outcome that tests whether the excess exposure had an effect on IQ.

In a review of epidemiological methodology on this topic, Paul et al. (2019) discuss the fact that the assessment of predictors of cognitive change (e.g. IQ decline) is most effective with longitudinal study designs that include repeated measurements of both exposures and outcomes. Only study designs #3 and #4 share these features.

It follows that study design #1 cannot determine if the BLL affected IQ or if IQ affected BLL. This study design cannot provide valid information on causal effects.

Study design #2 cannot determine how BLLs changed over the time period $t_0 - t_1$. It is unreasonable to assume that a BLL measured at time t_0 will not change in either direction after several years, given that lead exposures have been decreasing (in the population at large) and lead is not randomly distributed in the environment because it can be found in soil, water, food and beverages, and the air. In addition, study design #2 cannot determine how much and individual's IQ changed over the same time period $t_0 - t_1$. It is unreasonable to assume that the IQ measured at time t_1 is larger or smaller than the IQ (in the same individual) if it had been measured at time t_0.

Study design #2 cannot provide valid information on causal effects.

Study designs #3 and #4 can potentially test the hypothesis of interest, although study design #4 is a better test than study design #3.

Study design #3 has the problem discussed earlier in this report concerning measurements of IQ change.

Study design #4 is the only study design that approximates the structure of a randomized controlled trial, an important consideration in the evaluation of causality for individual epidemiological studies when environmental exposures are the issue (Dominici and Zigler, 2017).

Relevance of Study Design to the Lanphear et al. (2005) pooled analysis.

As described earlier in this report, the Lanphear et al. (2005) pooled analysis used data from 7 studies of blood lead levels and IQ. Importantly, the authors of the pooled analysis chose to use only one IQ test (measured somewhere between 4 years, 10 months, and 7 years). Similarly, the pooled analysis did not use sequential blood lead levels but rather single measures of blood lead, concurrent (i.e. closest to the IQ test), lifetime, maximum, and early childhood. The measure used primarily in the analyses and shown in Figures 1-4 (Lanphear et al. 2005, p. 898) was the concurrent measure. Simply put, the Lanphear et al. (2005) pooled analysis is an example of study design #1 above that cannot demonstrate causality. Indeed, the authors themselves recognize this deficiency; they write (Lanphear et al. 2005, p. 898):

> "The observational design of this study limits our ability to draw causal inferences."

Section 5.5 Application to the Population of Flint, Michigan

Furthermore, given that these studies have failed to provide adequate tests of the hypotheses that exposure to low levels of lead in childhood causes adverse neurodevelopmental outcomes, the application of the results of these studies to the population of children residing in Flint, Michigan who may have been exposed to a transient elevation in blood lead levels (due to the Flint Water Switch during several months in 2014 and 2015) cannot be scientifically appropriate. As discussed in detail

above, the Flint, Michigan population involved can be characterized as one in which many of the known causes of and risk factors for adverse neurodevelopmental outcomes are prevalent.

I conclude that the prevalence of causal factors and risk factors for adverse neurodevelopmental outcomes other than lead in women in Flint, Michigan during the years 2008-2015 is sufficient to explain the occurrence of those outcomes in children born during these years who were subsequently evaluated (e.g. in the years 2014 to the present). Put another way, given that the Lanphear et al. (2005) analysis fails to control for many of these factors—including but not limited to exposure to methylmercury, phthalates, PBDEs, and PCBs—it should not be used to make claims about the health of the children of Flint, Michigan who may or may not have been exposed to lead in water, depending upon their individual circumstances.

In the next part of this report, I critically examine the plaintiffs' experts' reports and opinions (including testimony).

PART SIX EVALUATION OF PLAINTIFFS' EXPERTS' REPORTS AND OPINIONS

Section 6.0 The Interdependence of the Plaintiffs' Experts' Claims

Before examining the validity and reliability of each of the plaintiffs' experts' claims, it is important to point out an important feature that connects them. Simply put, the reports of the plaintiffs' experts, Drs. Ducatman, Keating, Hu, Georgopoulos, and Lanphear are interdependent. Except for Dr. Lanphear, these plaintiff experts rely upon[5] the reports—and thus, opinions—of other plaintiffs' experts as shown in the following table:

	Relies on Ducatman?	Relies on Keating?	Relies on Hu?	Relies on Georgopoulos?	Relies on Lanphear?
Ducatman		Y	Y	N[a]	Y
Keating	Y		Y	N[a]	Y
Hu	N	N		Y	N
Georgopoulos	N	N	Y		N
Lanphear	N	N	N	N	

[a] Drs. Ducatman and Keating rely on Georgopoulos indirectly because they both rely on Hu.

What this table shows is that Dr. Ducatman relies on the reports and therefore the claims of Drs. Keating, Hu, and Lanphear (Ducatman Report, p. 12, Footnote 3). Dr. Keating relies upon the reports of Drs. Hu and Lanphear (Keating Report, p. 9); he also relies upon Dr. Ducatman's report (Keating Report, p. 15, 23, 24). Dr. Hu (Report, p. 14) relies upon the report of Dr. Georgopoulos who, in turn, relies upon the report of Dr. Hu (Georgopoulos Report, p. 6). Dr. Lanphear relies on no other plaintiffs' expert reports in his written report. There are other plaintiffs' experts involved in this litigation whose names could have appeared in this table. For my purposes here, I will limit my discussion to the experts named above.

The importance of these interconnections can be understood as follows. If the claims of an expert—say, "Expert H,"—are shown to be invalid and therefore unreliable, then those failures will likely if not assuredly affect the validity of the claims of the expert who relies upon the claims of "Expert H."

For example, Dr. Ducatman in his report relies upon the claims of Drs. Keating, Hu, and Lanphear. Dr. Hu's report and claims rely, in turn, upon the work of Dr. Georgopoulos. It follows that if Dr. Hu's claims—his opinions—are invalid then Dr. Ducatman's opinions, in turn, are if not completely invalid then certainly undermined in their validity. For example, given that Dr. Ducatman's program requires that a general causation relationship be established between exposure (lead) and outcome (decline in IQ) for all members of the "Minors Subclass," and given that he relies upon Dr. Hu's claims regarding that relationship, if Dr. Hu's claims are invalid then Dr. Ducatman's program has failed to meet its requirements. Another example involves the fact that Dr. Ducatman's proposed program relies upon Dr.

[5] I am limiting my analysis to explicit mentions of direct reliance in the plaintiffs' experts' reports. For example, Dr. Keating specifically mentions and refers to the reports of Drs. Ducatman, Hu, and Lanphear in his report. He does not mention the report of Dr. Georgopoulos.

Keating's claims regarding the interventions that are to be included in that program. I will show, however, that Dr. Keating provides no such interventions. It follows that Dr. Ducatman's program is lacking in substance and cannot be considered valid and reliable. In many ways, the plaintiffs' expert reports and claims are interconnected and interdependent.

To be sure, it is also possible—and in this case definite—that the claims of a plaintiffs' expert are invalid on their own, without the assistance of invalid or, for that matter, valid claims from other plaintiffs' experts.

In my discussion of various expert reports and claims, I will examine how these interconnections and interdependencies affect the validity and reliability of the opinions offered. I turn now to a careful examination of each of the plaintiffs' experts claims and the scientific approach that I take to assessing the validity and reliability of their claims.

Section 6.1 Approach to the Evaluation of the Opinions of Drs. Lanphear, Hu, Ducatman, and Keating

By an "evaluation" of the opinions of the plaintiffs' experts, Drs. Lanphear, Hu, Ducatman, and Keating in this matter, I mean an assessment of the extent to which their opinions are scientifically valid and reliable as well as ethically appropriate, given the arguments they make in their reports (and testimony when available), the scientific evidence that they use as support for their claims, and the methods (if any) they use to summarize and interpret that evidence.

The approach I take here is directly analogous to the process of peer review in scientific practice. There, a scientist's claims (and the evidence and methods used to make and support those claims) are subject to review and critique. The peer review process, as such, is familiar to and accepted by practicing scientists. It is an essential part of the practice of science. This process serves to increase the validity and reliability of the content—the results and interpretation of results—found in the scientific literature. A scientist's data, methods, and interpretations (e.g. causal claims) are subject to scrutiny by one's peers. Rejection of an author's claims (indeed, rejection of the manuscript as a whole) is not uncommon and occurs for many reasons, including when the methods are faulty (or nonexistent) and when the claims made are unjustified.

For example, the following are issues to consider when providing peer review (Weed, 2018):

1. whether the authors rely primarily upon method or primarily upon their subjective judgment to make claims
2. whether a method is described to be used in making claims (e.g. claims of causation)
3. whether the method described is one generally recognized in the scientific community and referenced there
4. whether the authors' description of that method is accurate (i.e., whether it reasonably conforms to the descriptions of that method in the published literature or misrepresents [i.e. deviates prominently from] those same descriptions)
5. whether that method is appropriate for the scientific question at hand and
6. whether the method selected by the author has been used appropriately to interpret results

Failure to meet any of these concerns would be grounds for determining that the expert's opinions or conclusions are invalid and/or unreliable. For example, failure to describe and use an established

methodology (see #1, #2 and #3 above) or misrepresenting/misapplying a methodology (see #4, #5 and #6) are particularly important considerations, calling into question the validity of the expert's claims. It follows that each plaintiffs' expert should provide an adequate (i.e. comprehensive and accurate) description of and justification for the methods he uses to make claims about causation or any other issue. Furthermore, the expert should show how the actual use of those methods in this matter justifies the claims made.

Additional Considerations in Peer Review

There are many other concerns in peer review, such as:

1. the extent to which the logic of an expert's argument is valid,
2. whether the expert has comprehensively searched for and cited the appropriate scientific literature, providing explicit criteria for including or excluding studies to be relied upon,
3. whether the expert's cited literature supports the facts, arguments, or conclusions claimed by the expert
4. whether the expert has provided an objective process for evaluating the quality of studies,
5. whether the expert has considered and accounted for the many forms of bias that can affect the valid interpretation of scientific evidence, and
6. whether the expert provides an approach for interpreting the evidence.
7. whether the expert has adequately considered the scientific and ethical issues associated with his claims
8. whether the expert provides the information they claim to provide

In this report, I will show that Dr. Lanphear did not provide adequate descriptions of the causal inference methods much less adequate citations to the peer-reviewed scientific literature where these methods have been discussed and codified. I will show that his general causation opinions are invalid and unreliable. In addition, I will show that Dr. Hu also fails to provide methodologic support for his claims and has unscientific—if not, absurd—claims about the nature of causation. Furthermore, I will show that the medical monitoring program proposed by Dr. Ducatman—that relies in part upon the claims of Dr. Keating—lacks sufficient information to be considered scientifically and ethically appropriate, and perhaps most importantly, is not an early detection program despite the fact that Dr. Ducatman claims it is one. In addition, Dr. Ducatman holds indefensible and unsupported beliefs about the nature of early detection.

In the following section, I describe these and additional errors committed by Drs. Lanphear, Hu, Ducatman, and Keating in their reports and related depositions.

Section 6.2 Dr. Bruce Lanphear

Dr. Lanphear has submitted a 16-page report in this matter, dated June 19, 2020. He was deposed on September 17 and 18, 2020. With a few relatively minor exceptions, this report is identical to a report Dr. Lanphear submitted in an earlier matter, i.e. Concerned Pastors for Social Action, et al., Plaintiffs, v. Nick A. Khouri, et al., Defendants (Case No. 16-10277, in the United States District Court for the Eastern District of Michigan, Southern Division), dated March 23, 2016. In addition, he was deposed in that earlier matter, on July 14, 2016.

Section 6.2.1 Dr. Lanphear's Purposes and Claims

In this section, I will describe Dr. Lanphear's stated purposes for participation in this litigation and the claims he makes regarding those stated purposes.

Dr. Lanphear's Report (June 19, 2020)

Regarding the purpose of his report and testimony, Dr. Lanphear describes the following (p. 1-2 @ 2):

> "I was retained to provide my expert opinions regarding the adverse health effects of lead poisoning, including the dangers to the health of children and adults in Flint resulting from their exposure to lead in Flint's public drinking water."

Furthermore, Dr. Lanphear writes (p. 1-2, @2):

> "my opinions concern whether there is any safe level of exposure to lead; whether children and fetuses are particularly vulnerable to the harmful effects of lead; whether the neuropsychological and developmental effects caused by childhood exposure to lead—including lowering IQs and academic achievement and increased risk of behavioral problems—are irreversible; whether the percentage of children in Flint with a blood lead level > 5µg/dL increased or 'spiked' from 2.4% to 4.9% as a result of the Flint Water Crisis; the impact of lead on hypertension, coronary heart disease, infertility and pregnancy outcomes in adults."

It is important to point out that Dr. Lanphear's only opinion that applies specifically to the "Minors Subclass" of the plaintiffs' class is the following:

> "Virtually all of the children and families in Flint were exposed to excess lead from drinking water during the period when chemicals were not adequately used for corrosion control. The excess lead resulted in substantial risk of harm to children and their families."

Dr. Lanphear provides no clue as to which adverse health outcomes he believes were experienced by every member of the "Minors Subclass." Rather, he only states that there was a "risk of harm" to children.

Dr. Lanphear's views contradict those of Dr. Hu, another plaintiffs' expert. As I will describe later in this report, Dr. Hu states that, regarding declines in IQ he alleges were experienced by the "Minors Subclass," there was no "risk" involved. These IQ declines, in his view, were as inexorable as if they were predicted from some natural or physical law absent any influence by random fluctuation, measurement error, other forms of systemic bias, or replacement by some other known cause of IQ declines in children. It follows that the 2 plaintiffs' experts—Drs. Lanphear and Dr. Hu—are at odds with one another. One, Dr. Lanphear admits that risk—also known as probability—is involved. The other, Dr. Hu, believes the opposite. I will have more to say about Dr. Hu's errors on this issue later in my report.

Dr. Lanphear's views also contradict Dr. Hu's views on the nature of the relationship between blood lead levels and IQ decline. Dr. Lanphear—the first author of the pooled analysis described earlier in this

report—supports the idea that this relationship is curvilinear—a curve—whereas Dr. Hu clearly supports the idea that the relationship is linear, a straight line. Both experts—as I will discuss in more detail—ignore the fact that the relationship is biased. This distinction is important in the context of the class certification issue because the Lanphear et al. (2005) curvilinear relationship involves differing amounts of IQ decline depending upon the starting blood lead level. In short, if a child has an initial (starting) blood lead level of 1 µg/dL, then the IQ decline is different than for a child with an initial (starting) blood lead level of 2 µg/dL.

Most relevant to my assessment of Dr. Lanphear's report are sections on his opinions (Section III, p. 5-12), facts and data considered (Section IV, p. 12), and a list of references (Section VIII, p. 13 and Exhibit 2). My focus here will be on Dr. Lanphear's approach (or lack thereof) for describing the relationship between exposure to lead and various health outcomes, as well as his assessment of the potential impacts of lead exposure from drinking water in Flint, Michigan.

Section 6.2.2 Dr. Lanphear's Failure to Define and Cite a Methodology

As described earlier in this report, the evaluation of scientific evidence for the purpose of determining causation (or the characteristics of a causal relationship) between exposure to lead (or any other factor) and health outcomes requires the explicit description of methodology. I include a description of these methods in Appendix D of this report. As noted there, and in a recent publication in the Journal of the National Cancer Institute, assessments of causality require that the author (or, in this case, the expert) perform a systematic review of the evidence (Weed, 2016; Weed, 2018).

Dr. Lanphear's report, however, does not include any explicit description of methodology. It is scientifically inappropriate if not irresponsible to make causal claims—as Dr. Lanphear does—in his report in the absence of a causal methodology.

I recognize that Dr. Lanphear claims in his report (p. 5 @10) that in arriving at his opinions he:

> "…relied upon reliable principles and methods which are generally accepted in the medical and scientific disciplines of pediatrics, toxicology, environmental health, and epidemiology."

However, Dr. Lanphear provides no citations to the peer-reviewed literature where those "principles and methods" are explicitly described and discussed. In the scientific and medical communities of pediatrics, toxicology, environmental health and epidemiology, a legitimate scientific practitioner does not simply state that they are using "reliable principles and methods" and then make claims without describing those principles and methods. This description should also provide explicit citations for where those "principles and methods" can be found in the peer-reviewed literature and then a demonstration of how the use of those "principles and methods" resulted in the claims made by that scientist. Dr. Lanphear provides none of this essential information.

Furthermore, Dr. Lanphear's comment (Report, p. 5 @9) that "all of the information set forth in this (his) report is based upon my education, personal knowledge, and experience, as well as my review of the documents cited in this report (see Exhibits 1 and 2)" lacks a sufficient basis for reliability.

Scientific claims—e.g. about causation—require explicit and careful description of the methods used to make those claims. In the absence of methodology, scientific claims about causality are no more than personal and thus subjective beliefs.

It follows that Dr. Lanphear's claims regarding causality are not based on methodology at all but rather on his personal and thus subjective beliefs.

Finally, Dr. Lanphear claims that his opinions arrived by reviewing documents. Once again, Dr. Lanphear has provided no explicit citation to the methodology used in his so-called "review." As I will show in detail immediately below, Dr. Lanphear has not performed a systematic review of the scientific evidence on lead and the several health outcomes he claims he will discuss. This is a fatal flaw. In the contemporary practice of pediatrics, toxicology, environmental health, and epidemiology, systematic reviews are considered necessary for making claims about causation. The scientific basis for this statement can be found in my section on methodology in Appendix D; there the reader will find extensive documentation in the scientific literature describing the need for systematic reviews when causality issues are being considered.

Indeed, as I have implied above, Dr. Lanphear provides no mention of any methodology for determining causation or for any of his statements. He provides no citations to the peer-reviewed literature where these methods have been described and discussed for decades. Beyond a statement about what Dr. Lanphear claims to be the purpose of his review, he does not provide the following:

1. explicit search terms and scientific literature databases searched,
2. explicit inclusion and exclusion criteria (for the studies to be reviewed),
3. explicit consideration of the so-called "grey" literature, i.e. unpublished reports, etc.
4. detailed descriptions (e.g. a table) of the characteristics of the included studies,
5. formal quality assessments of the included studies,
6. appropriate incorporation of the quality assessments in combining results, and
7. appropriate methods for combining results of the studies

In the absence of this methodology, the claims by Dr. Lanphear regarding the nature of lead exposure and the relationships between lead and health outcomes are based on his subjective views rather than on scientific method.

It might be countered that, after all, Dr. Lanphear has published (and cites in his report) a review entitled "The Impact of Toxins on the Developing Brain" in 2015 and therefore is justified in claiming that he has used (i.e. relied upon) a legitimate scientific publication for his many claims. A careful examination of this 2015 publication, however, reveals that it is also not a systematic review of the evidence on the "impact of toxins on the developing brain." It lacks all the same features listed above, that is, it lacks nearly all critical components of a systematic review of scientific evidence on topics such as the health effects of lead exposure. Furthermore, Dr. Lanphear's 2015 review does not include a discussion of adult outcomes, including but not limited to pregnancy outcomes and infertility.

Section 6.2.3 Dr. Lanphear's Claims: Inappropriate and Inadequate Citations to the Literature

Recognizing that Dr. Lanphear has failed to systematically review the evidence, I turn now to an assessment of the claims he makes and the citations to the scientific literature he provides in his June, 2020 report in this litigation matter. This is a key concern in a scientific assessment (i.e. peer review) of

any document that purports to be scientific as described above. When a scientist makes a statement or claim and provides a citation to the peer-reviewed literature (or to the so-called "grey" literature), there is an implicit expectation that the citation supports the statement or claim.

There are many examples in Dr. Lanphear's report where the citation he provides fails to adequately support the statements (and claims) he makes or where Dr. Lanphear provides no support at all from the scientific literature. The following examples are illustrative:

Claim #1: That "there is no safe level of exposure to lead" (Lanphear Report, p. 8 @ 23).

Dr. Lanphear provides no citations in support of this claim where he makes it. Later in his report, (see p. 9 @ 26) however, Dr. Lanphear states that "in 2012, the Centers for Disease Control and Prevention declared that there is no safe level of lead in children's blood" without citing the document from which this claim—whether true or not—can be found.

Dr. Lanphear is incorrect. The Centers for Disease Control and Prevention (CDC) did not state that there is no safe level of lead. Rather, the CDC (2012, p. 5) noted that "…no measurable level of blood lead is known to be without deleterious effects…" There is an important distinction to be made between not knowing whether a threshold exists and stating that there is no threshold level. As importantly, the CDC (2012, p. 6) report was not designed to assess "safety" per se but rather to determine a "reference value" for lead in children. Put another way, "safe" levels are, in fact, characteristic of the language of risk assessment, and the CDC (2012, p. 6) report specifically states that it is not intended as a risk assessment for lead. In any case, Dr. Lanphear has misrepresented the scientific evidence.

Claim #2: That "children are particularly vulnerable to the neurotoxic effects of lead because their brain is rapidly growing during fetal development and early childhood" (Lanphear Report, p. 8-9 @ 25).

Dr. Lanphear provides no citations in support of this claim.

Claim #3: That "IQ, academic achievement and behavioral problems (due to exposure to lead in children) are irreversible" (Lanphear Report, p. 9 @ 26).

Dr. Lanphear provides no citations in support of this claim but other plaintiffs' experts (e.g. Drs. Ducatman and Keating) claim that interventions can reverse these same problems.

Claim #4: That "adult lead exposure can result in increased blood pressure, or hypertension, and chronic kidney disease. Adult lead exposure has been associated with increased risk of cardiovascular problems, decreased cognitive function, and increased incidence of tremors" (Lanphear Report, p. 10 @ 29).

I will provide more details on why Dr. Lanphear has provided a scientifically inappropriate and insufficient citation for this claim regarding the proposed effects of lead on adults.

Claim #5: Dr. Lanphear claims (Report, p.2 @ 2) that "lead…(increases) the risk that (a) baby will be born too early or too small. Lead exposure has been associated with an increased incidence of miscarriages and delays in the time to achieve pregnancy."

Dr. Lanphear provides the following three citations for this set of claims that concern premature birth, births that are small for gestational age, miscarriages, and delays in the time to achieve pregnancy:

Edwards (2014), Zhu et al. (2010), and Borja-Aburto et al. (1999). It is relevant to point out that Dr. Lanphear also cites Taylor et al. (2015) in his reference list but does not cite it.

Claim #6: Dr. Lanphear claims (Report, p. that "virtually all" (but, in fact, not all) the children and families of Flint, Michigan were subject to a risk of harm.

> "Virtually all of the children and families in Flint were exposed to excess lead from drinking water during the period when chemicals were not adequately used for corrosion control. The excess lead resulted in substantial risk of harm to children and their families."

I have discussed the inadequacy of this claim above.

Section 6.2.4 Dr. Lanphear's Assessment of the Potential Effect of Lead Exposure on Health Outcomes in Adults

As noted earlier, Dr. Lanphear clearly states that he will provide expert opinions on "the impact of lead on hypertension, coronary heart disease, infertility and pregnancy outcomes in adults" (Lanphear Report, p. 2 @2).

Hypertension and Coronary Heart Disease: Dr. Lanphear's Failure to Provide Reliable and Valid Evidence

Regarding hypertension and coronary heart disease, Dr. Lanphear provides the following opinion (Report, p. 10 @29):

> "Adults exposed to lead can also experience adverse health impacts. Adult lead exposure can result in increased blood pressure, or hypertension, and chronic kidney disease. Adult lead exposure has been associated with increased risk of cardiovascular problems, decreased cognitive function, and increased incidence of tremors."

There is a single citation Dr. Lanphear provides at the end of these two sentences, i.e. Triantafyllidou and Edwards (2012).

The publication by Triantafyllidou and Edwards (2012), however, is not a review of the relationship between exposure to lead and health outcomes in adults, or for that matter, between exposure to lead and health outcomes in children. Rather, it is narrative and thus unsystematic review of the relationship between lead levels in water and lead levels in blood. Therefore, this 2012 publication is not an appropriate citation for Dr. Lanphear's claims. That said, the Triantafyllidou and Edwards (2012) publication provides a single table that Dr. Lanphear points to as providing support for his claim that "adult lead exposure can result in increased hypertension, chronic kidney disease, cardiovascular problems, decreased cognitive function, and increased incidence of tremors." I reproduce that table here, adapted from Triantafyllidou and Edwards (2012, Table 7, p. 1319):

Blood Lead Level (BLL) and Adverse Health Effects in Adults

BLL (µg/dL)	Adverse Health Effect
<10	Uncertain
>10	Hypertension
>20	Erythrocyte protorphyrin (+)

>30	Systolic blood pressure (+) Hearing (-)
>40	Nerve conduction (-) Infertility (men) Kidney failure
>50	Hemoglobin synthesis (-) Frank anemia Brain disorders
>100	Death

It is important to note that this table provides no information on chronic kidney disease, cardiovascular problems, decreased cognitive function, or increased incidence of tremors, the adverse health effects Dr. Lanphear claims not only are caused by (or associated with) exposure to lead but, according to him, find support in the Triantafyllidou and Edwards (2012) publication. Indeed, there is no mention of these health effects anywhere in this 2012 publication.

Clearly, Dr. Lanphear has provided profoundly insufficient evidence for his claim regarding the hypothesized health effects of lead exposure in adults. In addition, the only adverse health effect mentioned (although not discussed) in the Triantafyllidou and Edwards (2012) publication is hypertension which they note requires a blood lead level of greater than 10 µg/dL. Dr. Lanphear provides no evidence that any adult living in Flint, Michigan during the Flint Water Switch (FWS) had a blood lead level greater than 10 µg/dL that can be attributed to the FWS.

Pregnancy Outcomes and Infertility: Dr. Lanphear's Failure to Provide Reliable and Valid Evidence

Dr. Lanphear's claims regarding the hypothetical effects of lead exposure on pregnancy outcomes and infertility can be found in his report as follows (Report, p. 8 @ 24):

> "Lead can pass from a mother's lead store and blood to her unborn baby, increasing the risk that the baby will be born too early or too small. Lead has also been associated with an increased incidence of miscarriages and delays in time to achieve pregnancy."

The citations for these claims are Edwards (2014) and Zhu et al. (2010).

Dr. Lanphear continues (Report, p. 8 @ 24):

> "One case-control study showed that the odds of miscarriage nearly doubled for every 5 µg/dL increase in maternal blood lead concentration."

The citation for this claim is Borja-Aburto et al. (1999). Note also that Dr. Lanphear's list of references includes a paper by Taylor et al. (2015) not cited in his report.

Before dissecting the extent to which these four publications provide support for Dr. Lanphear's claims, it is important to examine the full body of evidence available to any legitimate scientist who carefully and systematically searches the scientific literature for studies, reviews, and commentary on the four outcomes Dr. Lanphear mentions: premature birth, low birth weight, miscarriages (more commonly called spontaneous abortions), and delays in fertility.

Put another way, the appropriate scientific approach to determining whether exposure to lead causes (or, for that matter, is associated with) any of these four pregnancy outcomes is to systematically search for the relevant literature, describe it, and then interpret that same evidence using methods commonly accepted in the scientific and medical communities. Clearly, Dr. Lanphear has not performed such a systematic assessment of the literature for these or any other outcomes mentioned by him in his report.

An Illustration of Dr. Lanphear's Failure to Systematically Assess the Literature on Lead and Preterm Birth (PTB), and Miscarriages

Consider the following examples. Dr. Lanphear's claim that exposure to lead "increases the risk" of premature birth (preterm birth) and babies born "too small" involves three of the four publications either cited by Dr. Lanphear [i.e. Edwards (2014) and Zhu et al. (2010)] or found in his reference list but not cited [i.e. Taylor et al. (2015)]. Because the phrase "too small" involves several possible outcomes, e.g. small for gestational age, birthweight, head circumference at birth, and length at birth, I will not examine all these outcomes in this report, until Dr. Lanphear has an opportunity to clarify his claim.

Spontaneous Abortions (i.e."Miscarriages")

As described in detail by Slama et al. (2013), "miscarriages" are typically referred to in the scientific literature as "spontaneous abortions," representing fetal loss between six (6) and twenty-two (22) weeks' gestation. Furthermore, "stillbirths" represent fetal loss after 22 weeks' gestation or, in some states, after 28 weeks' gestation. To be clear, "spontaneous abortions" and "stillbirths" represent two types of fetal loss. Perinatal mortality, in contrast, represents deaths occurring between 22 weeks' gestation through 1 week postpartum.

Dr. Lanphear's claim regarding the hypothetical effect of lead exposure on the occurrence of miscarriages (i.e. spontaneous abortions) involves three of the four papers cited in his report: Edwards (2014), Zhu et al. (2010), and Borja-Aburto et al. (1999). One might assume that these three papers represent the world's literature on the topic. But that would be an inappropriate assumption. Consider the following fact: Zhu et al. (2010) does not examine spontaneous abortion as an outcome. Thus, Dr. Lanphear's claim relies upon exactly two publications: Edwards (2014), an ecologic study that cannot demonstrate causality due to the ecologic fallacy and Borja-Aburto et al. (1999), to be discussed in more detail below.

In sum, Dr. Lanphear cites exactly one citation—a publication describing a case-control study—that he believes supports the claim that exposure to lead "has … been associated with an increased incidence of miscarriages…" (Lanphear Report, p. 8 @ 24). Dr. Lanphear writes that this "case-control study showed that the odds of miscarriage doubled for every 5 µg/dL increase in maternal blood lead concentration."

What is the available epidemiological evidence on the hypothetical relationship between exposure to lead and miscarriages (i.e. spontaneous abortions)?

In the past, **Hertz-Picciotto (2000), Antilla and Sallmen (1995)** as well as **Savitz et al. (1994)** reviewed the scientific literature on exposure to lead and other metals, as well as other occupational exposures on the occurrence of spontaneous abortions. Hertz-Picciotto (2000) concluded that "Although lead appears to play a role in excess fetal loss among those with high exposures, most studies of the risk from low-to-moderate-level exposures have provided little evidence for such reproductive toxicity. These studies, however, are far from definitive, due to numerous methodologic deficiencies." Antilla and

Sallmen (1995, p. 919) write that "exposure to lead may be associated with the risk of spontaneous abortion" based on only three studies with two of those studies having "methodological shortcomings such as misclassification of exposure, neglect of controlling of potential confounders or risk modifiers, or errors in the statistical analysis."

The **National Toxicology Program Report on the Health Effects of Low-Level Lead (2012)** examined the available literature on lead exposure (i.e. < 10 µg/dL) and spontaneous abortion, citing McMichael et al. (1986), Murphy et al. (1990), Borja-Aburto et al. (1999), Lamadrid-Figueroa et al. (2007), Yin et al. (2008), and Vigeh et al. (2010). This NTP Report concluded that there was "limited evidence that maternal blood Pb < 10 µg/dL is associated with spontaneous abortion" (2012, p. 107). This conclusion is based on 2 positive studies and 4 null studies, revealing a prominent inconsistency in the results, hence the label "limited evidence."

Dr. Lanphear's report (2020) cites only one of the 6 studies cited by the NTP Report [i.e. Borja-Aburto et al. 1999] revealing "cherry-picking" of the data. Furthermore, 4 of the studies cited by NTP (2012) reveal no association between maternal blood Pb and spontaneous abortion, i.e. McMichael et al. (1986), Murphy et al. (1990), Lamadrid-Figueroa et al. (2007), and Vigeh et al. (2010).

In order to provide a systematic assessment of the scientific evidence on exposure to lead and spontaneous abortion, I systematically searched the scientific literature and identified one additional study published since the 2012 NTP Report on this topic by Buck-Louis et al. (2017).

Buck-Louis et al. (2017) is the report of a cohort study—the "Life Study"—funded by the National Institute of Child Health and Human Development (NICHD), one of the National Institutes of Health. The study involved 501 couples recruited during 2005-2009. These couples had discontinued contraception with the intent of becoming pregnant. The final study population was 344 pregnant couples. In-person interviews, anthropometric measurements (e.g. BMI) and other information was collected. Blood was collected as well as trace element analyses including lead (Pb) and other metals (e.g. cadmium and mercury). The authors statistically analyzed time to pregnancy loss given exposure adjusting for age, BMI, history of prior loss, alcohol use, cigarette smoking, and creatinine clearance. Pregnancy loss was determined using an accurate pregnancy test, i.e. with a test with a false positive rate of 0 to 0.3%. Once the pregnancy test was positive, then pregnancy loss was determined by a later negative pregnancy test, clinical confirmation, or return of menses. The authors observed (Buck-Louis et al. 2017, Table 3, p. 72) that exposure to lead was not associated with pregnancy loss (all losses occurring prior to 22 weeks):

HR = 1.01 (95% CI: 0.82-1.25)

Buck-Louis et al. (2017) noted that the Borja-Abuerto (1999) study did not verify pregnancy loss. Indeed, Borja-Abuerto et al. (1999) relied on telephone interviews to assess pregnancy loss, thus exposing the study to bias.

It follows that of the several studies that have examined the relationship between exposure to lead and spontaneous abortion, the best designed study by Buck-Louis et al. (2017) revealed no association. It follows that the evidence on this topic is both inconsistent and insufficient to make any claim regarding the existence of an association much less causation. This conclusion is entirely consistent with the conclusion of the Agency for Toxic Substances and Disease Registry (ATSDR, 2020) which describes the evidence on lead and spontaneous abortion as inconsistent (ATSDR, 2020, p. 7).

Section 6.2.5 Summary of Dr. Lanphear's Claims

Dr. Lanphear's claims are not based on well-recognized and well-established methods. Rather, Dr. Lanphear's claims are based on hypothetical and untested assumptions and misrepresentations of existing literature. Dr. Lanphear's claims are not based on a systematic assessment of the literature. Some of his claims are not supported by any literature. For some of his claims, he provides a citation that, in fact, does not provide support.

For all the reasons and arguments above, I find Dr. Lanphear's opinions to be invalid and unreliable.

Section 6.3 Dr. Howard Hu

Dr. Hu has submitted a 45-page report in this matter and an accompanying list of references. In addition, Dr. Hu was deposed on October 12, 2020 and again on November 5, 2020.

Section 6.3.1 Dr. Hu's Purposes and Claims

In this section, I will describe Dr. Hu stated purposes for participation in this litigation and the claims he makes regarding those stated purposes.

Dr. Hu's Report

Regarding the purpose of his report and testimony, Dr. Hu describes the following (p. 1-2 @2):

> "I have been asked to evaluate whether such children (i.e. the putative subclass of children in the City of Flint who were exposed to, and ingested, Flint Water during the period of May 1, 2014 through January 5, 2016) suffered adverse health impacts as a result of the Flint Water Crisis and, if so, whether the existence of such adverse health impacts could be established using a methodology that is common to the subclass of children."

Regarding the nature of the "adverse health impacts" that Dr. Hu opines upon, these include decrements in IQ (Hu Report, p. 28 @23) and neurobehavioral disorders, specifically attention-related behavioral problems and antisocial behavioral problems, including attention-deficit/hyperactivity disorder (Hu Report, p. 42 and p. 43 @34).

Regarding Dr. Hu's claims—regarding causality—he states the following about the relationship between lead exposure and IQ and between lead exposure and, for want of a more precise term, neurobehavioral disorders:

See Hu Report (p. 27 @23) regarding lead and IQ:

1. "There is no known threshold of lead exposure below which such exposure is known to be safe."
2. "The greater the exposure to lead, the greater the adverse effects on health can be expected, i.e. the lead-adverse health effect relationship follows an incremental dose-incremental response relationship (monotonic sequence)."

Dr. Hu appears to believe that lead exposure causes decrements in IQ.

See Hu Report (p. 44 @34):

> "…similar to the blood lead level-IQ relationship, a threshold for lead's adverse impact on behavioral has not been seen even with respect to blood lead levels below 3 µg/dL, it is my opinion that lead exposure enough to establish an individual as a member of the injured class could very well be enough to have constituted a substantial contributor to a diagnosis of attention disorder or significant worsening of an existing attention disorder if it can be established that that the disorder or significant worsening of the disorder occurred 90 days or more after that individual's onset of lead exposure related to the Flint water crisis, and if the process of differential diagnosis and etiologic assessment was consistent with lead exposure as a substantial contributing factor."

Section 6.3.2 Dr. Hu's Contradictory Beliefs in How Exposure to Lead Affects IQ in Children

Based on his report, Dr. Hu believes that there is "no known threshold below which such exposure is known to be safe" as noted above.

However, he also believes that there is "no known threshold above the limit of…1 µg/dL" according to Dr. Hu's textbook account of the effects of lead (2018, p. 3298).

These are contradictory beliefs. Why? Because 1 µg/dL—according to Dr. Hu—is, in fact, a threshold. To put it another way, Dr. Hu is willing in litigation matters to ignore his views that he promotes for the medical community at large.

<u>Dr. Hu's Contradictory Beliefs in how Exposure to Lead affects Behavioral Disorders in Children</u>

There is another example of contradictory beliefs from Dr. Hu. Based on his report, Dr. Hu believes that low levels of lead exposure—i.e. below 3 µg/dL—causes attention deficit disorder (ADHD) and other behavioral disorders.

However, he also believes that this relationship requires much higher blood lead levels. In a most recent and classic (2018, p. 3298) textbook of internal medicine on this subject, Dr. Hu includes the following contradictory belief:

> "Subclinical exposures in children (BPb 25-60 µg/dL) are associated with deficits in…behavior and school performance."

To put it bluntly, in this legal matter, Dr. Hu is willing to claim that very low blood lead levels (< 3 µg/dL) cause behavioral disorders, but when he writes for the medical community at large, Dr. Hu claims that the blood lead levels needed to be "associated with" (and thus not necessarily cause) behavioral disorders are 8 to 20 times higher than the levels he believes are relevant for the members of the "Minors Subclass" in this litigation. Note that he had every opportunity in his textbook chapter account to include blood lead levels < 3 µg/dL but he did not do so.

Inconsistencies of this magnitude on the part of Dr. Hu cannot represent good scientific practice.

Section 6.3.3 Dr. Hu's Lack of a Valid and Reliable Methodology

Not only are Dr. Hu's views contradictory, they are also devoid of methodology.

Although Dr. Hu claims he uses methodologies, these methods are never described, much less applied. Examples follow.

Dr. Hu writes (p. 7 @8) that he employs "principles and methodology" that are "based on a review of an utilization of insights gained from peer-reviewed scientific literature that are relevant to the task at hand; the interpretation and utilization of publicly available data and data obtained in the discovery process that are relevant to the task at hand."

Note that he uses this "principles and methodology" language again in his Report on p. 13 @12.

He writes:

> "The task (i.e. the process of 'assessment of exposures and associated impacts') inherently requires a trans-disciplinary approach that integrates **principles and methods** related to water quality, chemistry, materials science, civil/environment engineering, exposure science, geostatistics, biological dosimetry, toxicokinetics, environmental epidemiology, and general environmental health."

At no point, however, does Dr. Hu define what principles and what methods he used to determine that lead exposure causes decrements in IQ or that lead exposure causes attention deficit disorder and other behavioral disorders.

Indeed, it appears that Dr. Hu's acceptance of a causal relationship between exposure to lead and the outcomes described above arise not from his own causal assessments but rather from his wholesale acceptance of claims made in the following documents:

1. Published studies and reviews.
2. Report of the Advisory Committee on Childhood Lead Poisoning Prevention of the Centers for Disease Control and Prevention (CDC), January 4, 2012.
3. The U.S. National Toxicology Program's (NTP) Report on the Health Effects of Low-Level Lead (June, 2012).
4. The European Food Safety Authority's (EFSA) Panel on Contaminants in the Food Chain (March 22, 2013).
5. The State of Michigan's LCR Workshop Presentation (undated).

Note that for the outcomes of attention-related behavioral problems and antisocial behavioral problems, Dr. Hu also relies upon a systematic review and meta-analysis by He et al. (2019) and a systematic review by Donzelli et al. (2019). See Hu Report (p. 43 @34).

What Dr. Hu cites from these documents is as follows:

1. That these publications "relate blood lead levels, going down to a level of 1 µg/dL, with decrements in IQ in children."

2. That "because no measurable level of blood lead is known to be without deleterious effects, and because once engendered, the effects appear to be irreversible in the absence of any other interventions…"
3. That "there is sufficient evidence that blood Pb levels < 5 µg/dL are associated with decreases in broad based and specific indices of cognitive function and an increase in attention-related behavioral problems and antisocial behavioral problems"
4. That "there is no evidence for a threshold for critical lead-induced effects"
5. That "no level of lead in the blood is safe."

It is important to point out that the word "cause" or "causality" does not appear in any of the quotes Dr. Hu includes in his report regarding these sources. It would be reasonable to infer that causality is assumed in some of these statements given that the word "effects" appears in #2 (the CDC 2012 Report) and #4 (the EFSA 2013 Report). On the other hand, the NTP (2012) Report uses the language of "association" rather than causation.

It follows that while Dr. Hu may believe that lead exposure causes decrements in IQ and causes attention-deficit hyperactivity disorder (and presumably its symptoms) as well as antisocial behavioral problems based on the documents cited, he has provided no independent analysis of the evidence and no description much less application of the methodology he used to assess causality.

Dr. Hu mentions and relies upon the exposure assessment methods found in other plaintiffs' experts' reports. This is an important point. Dr. Hu's report basically cedes all discussion of exposure assessment methodology to the reports (and therefore, opinions) of Drs. Russell, Weisel, Georgopoulos, and Goovaerts. But Dr. Hu's own methodology regarding causality is also what is at stake here.

Put another way, exposure assessment methods for the Flint population are insufficient for the purpose of determining causation at the population level, i.e. to determine whether and to what extent lead causes harm. This concern applies to the relationship between lead exposure and IQ as well as the relationship between lead exposure and behavioral disorders.

Given the brief discussion above, Dr. Hu has a serious if not fatal problem. His report has no methods section. He provides no mention much less a description of his methodology for determining causation.

As such his report is basically a subjective essay that cannot be considered objective. In common scientific parlance, it is a commentary.

As I have discussed at length in my report, the scientific community relies upon methodology for claims of causation. At the center of these methods is the systematic review. However, Dr. Hu's report does not include any of the components of a systematic review. The components are as follows:

1. explicit search terms and scientific literature databases searched,
2. explicit inclusion and exclusion criteria (for the studies to be reviewed),
3. explicit consideration of the so-called "grey" literature, i.e. unpublished reports, etc.
4. detailed descriptions (e.g. a table) of the characteristics of the included studies,
5. formal quality assessments of the included studies,
6. appropriate incorporation of the quality assessments in combining results, and
7. appropriate methods for combining results of the studies

None of these can be found in Dr. Hu's report. It follows that Dr. Hu's report is not systematic and his opinions regarding causality are invalid and unreliable on that basis alone, ignoring his inconsistencies and logical missteps.

Section 6.3.4 Dr. Hu's Claim that All the Children in the "Minors Subclass" were Harmed

According to Dr. Hu, every Flint, Michigan child who is a member of the "Minors Subclass" as defined in his report and in the accompanying complaint was injured by lead exposure that, in turn, was caused by the "Flint Water Switch (FWS)."

See Dr. Hu's report (p. 9-10), where he writes:

> "…the exposure—(i.e. as defined in his report and modeled by other plaintiffs' experts)—is of sufficient duration and magnitude such that each child will have sustained non-negligible impairment of their neurobehavioral development."

The minimum duration of lead exposure, according to Dr. Hu, that causes "impairment" in each and every child in the "Minors Subclass" is 14 days within a 90-day period during April 25, 2014 through October 16, 2015. See also the Class Plaintiffs' Motion for Class Certification (p. 2).

Dr. Hu then claims that the impact on the IQ of every child in the "Minors Subclass" can be determined using the regression function in Lanphear et al. (2005, 2019) along with the estimated blood lead level "specific for the contribution of the Flint water crisis" (Hu Report, p. 23). However, as noted earlier, Dr. Hu would use a linear rather than a curvilinear regression function. The Lanphear et al. (2005) publication clearly provides a curvilinear function for the relationship between blood lead levels and IQ decline.

Dr. Hu's approach means, therefore, that the best way to determine the IQ deficit of any child who is a member of the "Minors Subclass" and caused by the FWS is to assume that each such child sits, in essence, precisely on the Lanphear et al. (2005, 2019) regression curve. It follows that if such a child had an increase of 2 µg/dL of blood lead (presumably due to the Flint Water Switch), then their IQ deficit would be (2 X 0.51 µg/dL) or 1.02 IQ points. Similarly, if the child's blood lead elevation was 1 µg/dL then that individual would have lost 0.51 IQ points. Finally, if the child's blood lead elevation was 0.1 µg/dL, then that individual would have lost exactly 0.051 IQ points. It is important to point out that Dr. Hu believes that IQ deficits of 1.02 or 0.51 IQ points are not clinically insignificant in direct contrast to the published scientific literature as discussed earlier in this report. Recall that Haier (2014) concludes that even small statistically significant changes in test scores are not sufficient proof that general intelligence has changed. Breslau et al. (2001) write that even a 5-point IQ change in an individual is considered clinically insignificant (Breslau et al. 2001, p. 716).

Dr. Hu, on the other hand, writes that "although some critics have questioned the importance of small decrements in the IQs of individual children," he believes that "any detectable effect occurring from a widespread exposure is a cause for concern." (Hu Report, p. 26). To be clear, Dr. Hu does not provide any legitimate defense of this claim. Rather, he argues that small decrements in individual children's IQ can have an impact on the population-at-large. But that argument ignores the issue at hand which is whether IQ deficits of 1.02 or 0.51 (or, for that matter, 0.051) points are clinically insignificant at the individual level. The scientific community believes they are as I have described earlier in this report. Dr. Hu disagrees based solely on his personal subjective opinion and on an argument that avoids the issue.

This is just one more example of Dr. Hu's unscientific and, frankly, biased approach to the legitimate scientific questions at the center of this legal matter.

Dr. Hu has failed to provide scientific evidence that very small changes in IQ are clinically significant. Any member of the "Minors Subclass" who has experienced very small IQ deficits cannot be considered "harmed," in part because these deficits are not clinically significant and because of the likely possibility that any decline was caused by a variety of other well-established factors has not been ruled out by Dr. Hu. For a description of these factors, see Part Two of this report.

Section 6.3.5 Another Methodological Error in Dr. Hu's Report

In his report, Dr. Hu writes that he believes that the lead-IQ relationship defined by the Lanphear et al. (2005) curve "is likely higher" because "random error…is well known to bias estimates of effect towards the null" (Hu Report, p. 26-7).

This claim by Dr. Hu is scientifically nonsensical. Random error, by definition, is not a systematic error—a bias—in one direction or another. If that were the case, then random error would not be random. It would, in contrast, be considered a bias, i.e. a systematic error. But the scientific literature makes it clear that random error is not systematic error. Random error is not a bias. Consider these examples from a well-regarded textbook of epidemiology and from the Dictionary of Epidemiology:

> "Unlike bias and confounding, random errors are considered unsystematic because they arise from an unforeseeable and unpredictable process" (Aschengrau and Seage, 2003, p. 300).
>
> "Two broad kinds of error can occur in studies in the health, life, and social sciences (Porta, 2008, p. 85):
> 1. Random error: the portion of variation in a measurement that has no apparent connection to any other measurement or variable, generally regarded as due to chance.
> 2. Systematic error: error that is consistently wrong in a particular direction."

Dr. Hu is wrong. Random error does not bias estimates of effect towards the null.

Section 6.3.6 Dr. Hu's Failure to Incorporate Blood Lead Data on the Flint Children and Other Problems

Although blood lead levels are available for some children who presumably are members of the "Minors Subclass," Dr. Hu ignores this information. Nor does he have any actual measurement of the change in blood lead level for any child who is a member of the "Minors Subclass." Similarly, Dr. Hu has no actual measurement of the proposed change (i.e. deficit) in IQ of any child who is a member of the "Minors Subclass," i.e. a pre- and post-FWS IQ measurement. The lack of empirical testing of the regression model used by Dr. Hu is a fatal flaw from a scientific perspective. Lack of testability makes claims that purport to be scientific unscientific. Furthermore, Dr. Hu made no effort to obtain information on IQ testing on any of the plaintiffs in the "Minors Subclass."

Remarkably, Dr. Hu writes that measuring pre- and post-FWS IQ measurements for the children in the "Minors Subclass" would not provide reliable information. He writes (Hu Declaration, p. 33):

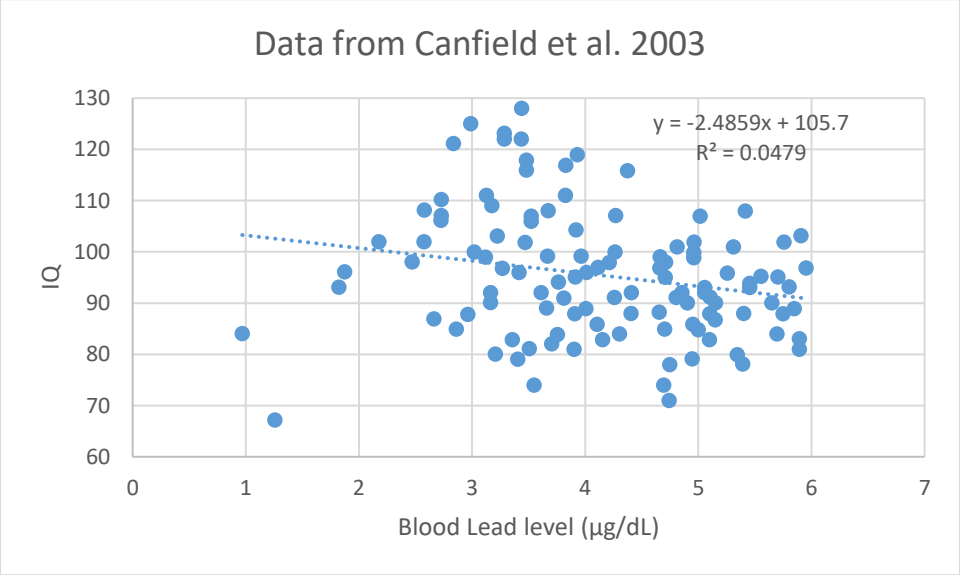
> "…such pre- and post-testing may not be able to reliably distinguish true changes in IQ from the random noise that typically occurs when IQ tests are repeated in the same individual."

If pre- and post-testing cannot be distinguished from random noise, then it follows that there is no way to test the validity and reliability of Dr. Hu's approach to determining the IQ change for each individual child who is a member of the "Minors Subclass." To put it another way, even if Dr. Hu had information on the change in IQ for the members of the "Minors Subclass," his testimony is that it would be scientifically unreliable. It is hard to imagine a more blatantly unscientific approach than what Dr. Hu claims. The bottom line here is that Dr. Hu is arguing that a completely hypothetical model—that uses no empirical data from the members of the "Minors Subclass"—is a better scientific approach than obtaining empirical data from those same members.

In essence, Dr. Hu's approach to determining the proposed IQ deficit for every member of the "Minors Subclass" is based on speculative (hypothetical) lead exposure data and the Lanphear et al. (2005, 2019) regression analysis.

Serious Problems with Using the Lanphear et al. (2005) Study to Predict Changes in IQ

Turning to the Lanphear et al. (2005, 2019) publication, it is important to point out that this pooled analysis used data (i.e. information) from 7 different epidemiological studies. In each such study, the authors estimated a curve (through regression) of the relationship between blood lead levels and IQ. A good illustration of the relationship between the actual measurements on the study participants and the calculated regression curve can be found in Canfield et al. (2003), one of the studies incorporated into the Lanphear et al. (2005) pooled study. See Figure 2 (p. 1525) of that publication and shown below:



What this figure reveals is that few—perhaps only a handful—of the individual data points of the study participants actually fall on the calculated regression curve. The data for most—nearly all—study participants does not fall on the calculated regression curve but rather either above or below the regression line.

The same situation exists for the study of lead levels and IQ published by Wasserman et al. (1997). See Figure 2, p. 960. Of the 309 children examined in this study, at most 5 (approximately 1.6%) of the data points fall on the regression line calculated.

It follows that the same phenomenon exists for the Lanphear (2005) regression analysis, namely, that the data from relatively few study participants fall on the regression line. Indeed, a careful examination of Figure 2 on p. 898 of the Lanphear et al. (2005) publication reveals that none of the curves generated in the 7 studies used in that pooled analysis actually fall on the regression line created by Lanphear et al. (2005). It follows that the measured data from few if any of the study participants in all 7 studies fall on the Lanphear et al. (2005) regression line.

This situation contrasts sharply with the premise of Dr. Hu's claim that every Flint, Michigan child's IQ deficit is determined using the precise information of the Lanphear et al. (2005) regression line.

It follows that the Lanphear et al. (2005) regression line does not actually predict the relationship between blood lead levels and IQ at the individual level given that the data from very few individuals – the study participants—actually falls on that regression line. As importantly, there is no information on the change in blood lead level much less the change in IQ for each plaintiff in the "Minors Subclass." Without that information, it is impossible to predict, much less know, whether each plaintiff was harmed as a result of the Flint Water Switch.

<u>On Dr. Hu's Belief that the Relationship Between Lead Levels and IQ is "Not Probabilistic"</u>

In this section, I will describe a fundamental flaw in Dr. Hu's testimony in this legal matter related to the discussion of the Lanphear et al. (2005) study above. The issue, as discussed below, involves Dr. Hu's belief that the relationship between lead levels and IQ is not a probabilistic one, meaning that there is no room in this relationship for any error, either systematic or random. As I will show, Dr. Hu testifies that "risk" has nothing to do with this relationship. The term for this type of relationship in science is "deterministic." Nothing could be further from the truth. Dr. Hu's views are so contrary to current scientific thinking as to be considered something akin to "junk science." My explanation follows.

According to Dr. Hu in his deposition testimony on October 12, 2020 (p. 95, lines 3-6), every child in the "Minors Subclass" in this litigation (regardless of whether there is evidence to the contrary) had an IQ deficit due to (i.e. caused by) the Flint Water Switch.

> "…all children who had significant elevations in blood lead, whether documented or extrapolated, suffered injury in the form of declining IQ."

Recall however that Dr. Hu ignores the "documented" blood lead levels measured in the children in the "Minors Subclass." Rather, Dr. Hu uses only the "extrapolated"—meaning assumed without regard for actual data—blood lead level from which IQ deficits are predicted.

Dr. Hu repeats this claim during the same exchange in his deposition (p. 95, lines 19-22):

> "Another way of saying that is that there's no evidence I'm aware of that indicates that a child who's experienced elevations in blood lead will have no effects."

Note that the phrase "experienced elevations in blood lead" is not an empirical fact but rather a hypothetical belief without regard for any measurement of blood lead in any child who is a member of the "Minors Subclass."

Dr. Hu's claim is both extraordinary and scientifically untenable. Simply put, Dr. Hu believes that the relationship between blood lead levels and IQ deficits is not probabilistic and therefore is deterministic, i.e. any (hypothetical) blood lead level change will definitely and inexorably lead to an IQ deficit. Indeed, Dr. Hu stated this idea explicitly during his deposition testifying about the relationship between lead levels and IQ (p. 95, lines 7-17):

> "That's not a risk. That is simply a direct connection based on the science. Risk, in my opinion, is a term better reserved for discrete outcomes, particularly diagnostic outcomes like cancer or in the case of lead exposure, something like attention deficit disorder. And that's when we start talking about probabilities. But in terms of elevations in blood lead and IQ, I don't think that's a probabilistic relationship."

What Dr. Hu is proposing here is that there is, in essence, a "natural physical law" between the two measures, blood lead levels and IQ. There is, in his view, no room for any error, bias, or random fluctuation. According to Dr. Hu, "risk" and "probability" are not involved.

This is an extraordinary and, frankly, ludicrous claim. There is no precedent in the literature on epidemiologic methodology for predicting a quantitative outcome (i.e. IQ) from an assumed (estimated) value—the blood lead level generated from Lanphear et al. (2005)—that can be made without any concern for error, whether systematic error (i.e. bias) or random error (chance). According to Dr. Hu, and assumed by all the plaintiffs' experts, once a blood lead level increase is estimated from the Georgopoulos model for a member of the "Minors Subclass," that same member will have experienced a deficit in IQ. There are no exceptions and no allowance for variability.

Dr. Hu's "no risk" claim goes against everything epidemiologists know about exposure-disease relationships. These relationships have, at their core, a fundamental notion that causality itself is a probabilistic idea (Parascandola and Weed, 2001). Dr. Hu's claim—made in deposition—that "risk" and therefore "probability" are not involved in this situation is, to be frank, absurd. There is no basis in the epidemiologic literature to claim that an exposure like lead that purports to represent a predicted a change in IQ can be valid without taking into consideration chance, bias, and confounding. In addition, it is important to remember that the calculation of an IQ decline for any individual member of the "Minors Subclass" is hypothetical, devoid of any information about that individual. Hypothetical, indeed, all epidemiological data applied to people not directly involved in the study must involve considerations of chance, bias, and confounding (Boffetta et al. 2020).

Dr. Hu's claim involves an expression of precision that, in this circumstance, is not only inappropriate but also cannot be tested with empirical data. Indeed, Dr. Hu notes, as I described earlier, that he believes actual measurements of IQ in children are imprecise and replete with random noise, i.e. random fluctuations in test values. It follows that Dr. Hu's application of the Lanphear et al. (2005) model to

determine (not estimate) changes in IQ cannot be tested empirically. Lack of empirical testability reduces Dr. Hu's claims to speculative unscientific statements.

To make matters worse, Dr. Hu's testimony contradicts his claim that the relationship between blood lead and IQ is not a probabilistic relationship and therefore is a deterministic relationship. As noted above, he testifies that "…all children who had significant elevations in blood lead, whether documented or extrapolated, suffered injury in the form of declining IQ." The key phrase here is "significant elevations."

But Dr. Hu's claim that the relationship is deterministic means that any blood lead elevation—whether "significant" or not—would cause an injury. Simply put, Dr. Hu cannot have it both ways. He cannot, on the one hand, exclude some children from his deterministic relationship because their blood lead elevation was not "significant" enough to change IQ, and, at the same time, maintain that any elevation in blood lead level leads inexorably to an IQ deficit. Either all blood lead levels lead to IQ change or some blood lead levels—due to systematic errors or random errors—do not lead to IQ change. Given that it is well established that, at most, blood lead levels contribute approximately 4% of the variability in IQ levels, there is plenty of room for systematic and/or random error in estimating the potential effect of blood lead on IQ at the individual or group level.

As described above, few if any of the study participants in the 7 studies used by Lanphear et al. (2005) fall on the Lanphear et al. (2005) regression line. But for this line to represent the "truth," i.e. the deterministic law between blood lead levels and IQ, every study participant would need to fall precisely on that same line. They do not. It follows that the relationship between blood lead levels and IQ is not "deterministic" but is, in fact, probabilistic.

The authors of the Lanphear et al. (2005) explicitly state that their results do not permit the conclusion that there is a causal relationship between blood lead levels and IQ, much less a deterministic relationship that is unaffected by random error or systematic error (bias). They write (Lanphear et al. 2005, p. 898):

> "The observational design of this study limits our ability to draw causal inferences."

Section 6.3.7 Summary of Dr. Hu's Report and Testimony

Dr. Hu's claims are not based on well-recognized and well-established methods. Rather, Dr. Hu's claims are based on hypothetical and untested assumptions and misrepresentations of existing literature. Dr. Hu's claims are not based on a systematic assessment of the literature but rather a highly selective approach that ignores a body of evidence on alternative models for the lead-IQ relationship and an even larger body of evidence on alternative causes for IQ declines in children. Dr. Hu's claims are contradictory, illogical and contrary to well-established scientific maxims; one prominent example is his claim that the relationship between lead exposure and IQ is "not probabilistic."

For all these reasons, Dr. Hu's opinions are invalid and unreliable.

Section 6.4 Dr. Alan Ducatman

Dr. Ducatman has submitted a 33-page report in this matter, dated June 28, 2020. He was deposed on November 12, 2020 and again on December 7, 2020. The purpose of Dr. Ducatman's report is to

propose an early detection—i.e. medical monitoring—program for the members of the "Minors Subclass."[6]

Section 6.4.1 Basic Principles and Practice of Medical Monitoring

With Dr. Ducatman's opinions in mind, I will discuss the appropriateness of medical monitoring, i.e. under what circumstances, if any, medical monitoring should be provided to the plaintiffs in this case for the purpose of early detection and treatment of conditions allegedly associated with exposure to lead in the children of Flint, Michigan.

Some background is necessary, namely, the basic scientific and ethical principles that are used in current medical (and public health) practice to determine the appropriateness of medical monitoring. I will then apply these principles to the early detection and treatment of the conditions and diseases at issue in this case: declines in IQ and behavioral disorders (attention deficit disorder and its related symptomatology), as these relate to lead exposure in the children of Flint, Michigan who are members of the "Minors Subclass."

<u>The Purpose of Medical Monitoring</u>

The purpose of medical monitoring is the early detection and treatment of diseases. There is no dispute about this fact. Textbooks (Rosenstock, 2005) and the Agency for Toxic Substances and Disease Registry (ATSDR, 1995) notices in the Federal Register—relied upon by Dr. Ducatman—make plain this fact. For example, from a textbook of clinical occupational and environmental medicine (Rosenstock, 2005):

> "A medical surveillance program must be predicated on the early detection of the effects of an exposure, at a point where intervention can prevent disease or disability."

An ATSDR document (Federal Register, July 28, 1995) also unequivocably equates medical monitoring with screening of asymptomatic individuals to detect early disease:

> "…The monitoring aspect of a health surveillance program consists of the periodic medical testing to screen individuals who are at increased risk of disease. **Monitoring serves to identify those individuals with an unrecognized adverse health effect. This is consistent with the definition of screening as 'the presumptive identification of unrecognized disease or defect by the application of tests, examinations or other procedures which can be applied rapidly. Screening tests sort out apparently well persons who probably have a disease from those who probably do not.** A screening test is not intended to be diagnostic. Persons with positive or suspicious findings must be referred to their physicians for diagnosis and necessary treatment…" (emphasis added)

[6] Also relevant to Dr. Ducatman's opinions in this matter are his expert opinions in another recent matter, specifically, a case involving exposure to PFOA, a chemical that allegedly contaminated drinking water. See also Ducatman Deposition Vol. I, p. 61. I will refer to Dr. Ducatman's report in the PFOA matter (dated, March 27, 2020) as needed given that his opinions on medical monitoring in that case are relevant to his opinions on medical monitoring in the Veolia case.

From these accounts, both of which make it clear that medical monitoring is about early detection (and treatment if appropriate), it follows that the exposure of interest must have had an effect established—i.e. a causal effect—and that the intervention(s) provided must be effective in preventing the outcomes of interest or reducing the impacts of those outcomes. Intervening on the effects of exposures that do not subsequently lead to (i.e. cause) disease cannot prevent disease. In any case, the undisputed aim of medical monitoring is the early detection of disease and intervention (treatment).

The Scientific Basis for Screening (Early Detection and Treatment)

The scientific principles that govern screening are well-recognized. The tests used must have sufficient specificity and sensitivity to be considered effective. The treatment (intervention) offered to those who test positive must be effective (Mant and Fowler, 1990; Holland, 1993). "Effective" in the screening context means that the intervention has been shown to reduce overall mortality or morbidity in well-designed clinical trials or (when trials are not possible) large bodies of relevant epidemiological evidence that have rigorously tested the hypothesis that the intervention does, in fact, reduce mortality or morbidity in populations. In these studies, the populations are tested (with the screening test) and those individuals who test positive are provided treatment after a diagnostic test confirms the finding. Thus, the tests and the treatments (interventions) are closely linked.

Expert groups, such as the U.S. Preventive Services Task Force, regularly make recommendations for screening based on a well-accepted and rigorous examination of the evidence in "evidence-based reviews" (Guirguis-Blake et al., 2007). Despite the popular (and mistaken) view that any test that finds disease early is necessarily beneficial (Sackett, 1987), screening for disease is not justified by high yields of disease (from the test), early stage at diagnosis, or even improved survival in the absence of control groups (Mant and Fowler, 1990).

It is relevant and important to point out that the U.S. Preventive Services Task Force does not recommend any medical monitoring (screening)—early detection and treatment—program for children or pregnant women with elevated blood lead levels regardless of the absolute value of that blood lead level. The USPSTF (2019, p. 1502) statement follows:

> "**The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of screening for elevated blood lead levels in asymptomatic children.** The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of screening for elevated blood lead levels in asymptomatic pregnant persons." (emphasis added)

Dr. Ducatman ignored this key organizational decision by the U.S. Preventive Services Task Force that clearly contradicts his view on the need for and structure of a program for medical monitoring the "Minors Subclass."

The Ethical Basis for Screening (Early Detection and Treatment)

Screening programs are guided not only by scientific but also ethical principles. The ethical principles that govern screening are also well-recognized and go well beyond informed consent (Marshall, 1996; Schwartz and Meslin, 2008) to include respect for autonomy, beneficence, and nonmaleficence. An ethical imperative (i.e. a duty or obligation) of anyone offering a screening program is that there must be a benefit to the individual who is tested. Secondly, the benefit must outweigh any harm (Mant and Fowler, 1990). Screening—in general—has been demonstrated to cause harm (Stoate, 1989). In

addition, participants in a screening program must be provided with adequate information regarding the risks and benefits of screening (Edwards and Hall, 1992; Schwartz and Meslin, 2008). Identifying disease early does not, by itself, provide the individual a benefit unless the treatment offered reduces morbidity and/or mortality as discussed above. In the absence of a demonstrated benefit, screening identifies an individual as "ill" who was previously "well" but the natural history of the disease process in that same individual will not be changed (because the intervention has no demonstrated benefit). Screening for disease in the absence of a well-tested efficacious intervention is ethically inappropriate, causing more harm than benefit. As Charlton has noted (1992, p. 521):

> "If there is no strong evidence of benefit the doctor should leave the patient alone. To do otherwise is, it seems to me, unethical and against the patient's interest."

Summary

From this brief description and discussion of the scientific and ethical requirements for screening programs, it will be important to assess the extent to which Dr. Ducatman provides adequate support for the program he proposes for the "Minors Subclass." To be specific, it will be important to examine whether Dr. Ducatman can provide evidence that specific and sensitive tests for the early detection of adverse outcomes allegedly due to lead exposure not only exist but have been tested for efficacy in clinical and/or epidemiological studies. It will also be important to examine whether Dr. Ducatman provides evidence that early intervention on children exposed to lead at blood lead levels experienced by the members of the "Minors Subclass" has resulted in improvements in reversing IQ decline or the occurrence of behavioral disorders. Finally, it will be important to examine whether Dr. Ducatman includes provisions for informed consent in the medical monitoring program as well as explicit consideration of the risks and benefits of the tests and interventions proposed.

In the absence of this information, Dr. Ducatman's proposed program is scientifically and ethically bankrupt.

Section 6.4.2 Dr. Ducatman's Purposes and Claims

In this section, I will describe in more detail Dr. Ducatman's stated purposes for participation in this litigation, the claims he makes regarding those stated purposes, and the support he attempts to provide for advocating for such a program.

Dr. Ducatman states the following (Report, p. 7):

He is going to "provide this report in support of ongoing, coordinated, diagnostic, evaluation, programmatic relief to mitigate the health damages including human performance damages to a clearly defined class of lead-exposed children in Flint, Michigan." (For confirmation, see Ducatman Deposition Vol. I, p. 46-7).

Dr. Ducatman relies upon three organizations for support: (1) the Pediatric Environmental Health Specialty Units (PEHSU) and American Academy of Pediatrics (AAP) and their guidelines, (2) the Centers for Disease Control's (CDC) and their document entitled, "Educational Interventions for Children Affected by Lead," and (3) the Agency for Toxic Substances and Disease Registry's (ATSDR's) and their medical monitoring program guidance document (1995).

Of these documents and as discussed above, the ATSDR guidance, makes clear that the program Dr. Ducatman proposes is, at its center, should be designed and undertaken as a medical monitoring program with early detection and intervention at its core. The concept of "early detection" is key.

Summary of the Recommendations in Dr. Ducatman's Primary Support Documents

PEHSU and American Academy of Pediatrics (www.pehsu.net)

Dr. Ducatman relies upon a document produced by PEHSU and the American Academy of Pediatrics (PEHSU-AAP) and entitled, "Recommendations on Medical Management of Childhood Lead Exposure and Poisoning." That document explicitly states that for children whose blood lead level is < 5 µg/dL there are no recommendations for "mitigation." These organizations recommend only that these children have repeat blood lead levels in 6-12 months and for medical professionals to "perform routine health maintenance" procedures such as assessment of nutrition, physical and mental development and provide "anticipatory guidance on common sources of environmental lead exposure."

In sum, the PEHSU-AAP recommendations do not recommend any developmental testing or education or nutritional intervention for children whose blood lead levels are less than 5 µg/dL. For children whose blood levels are between 5 µg/dL and 14 µg/dL, the PEHSU-AAP recommendations include nutritional counseling and "structured developmental screening evaluations" at child health maintenance visits. The PEHSU-AAP recommendations do not include any mention of interventions aimed at the mitigation of adverse intellectual or behavioral outcomes.

It is important to point out that Dr. Ducatman's reliance on the PEHSU-AAP document means that any members of the "Minors Subclass" whose blood lead levels were less than 5 µg/dL during the few months that Veolia was involved in the Flint Water Switch should not be given developmental testing or educational or nutritional intervention. If all members of the "Minors Subclass" had blood lead levels less than 5 µg/dL, then none should be tested or intervened upon.

Finally, Dr. Ducatman fails to point out that the PEHSU-AAP guidelines differ depending upon blood lead levels. Dr. Ducatman fails to consider the extent to which the children in the "Minors Subclass" fit into the various categories described in the PEHSU-AAP document.

The Centers for Disease Control and Prevention (CDC)

Dr. Ducatman also relies upon the CDC's document entitled, "Educational Interventions for Children Affected by Lead."

The CDC document (2015) specifically states that a specific evidence-base for educational interventions in children exposed to lead does not exist. The authors write (2015, p. vii) the following:

> **"There are no studies that specifically examine the impact of early childhood educational interventions on cognitive or behavioral outcomes for children who have been exposed to lead."** (emphasis added)

While Dr. Ducatman provides additional information from the CDC document, he fails to point out that none of the recommended interventions have been tested for children exposed to lead.

From these two documents—one from the PEHSU and the American Academy of Pediatrics and the other from the CDC—there are no (zero) evidence-based recommendations to provide interventions to children with blood lead levels between 0 µg/dL and 14 µg/dL. Dr. Ducatman's proposed program is devoid of evidence supporting interventions specifically designed for children exposed to lead.

Agency for Toxic Substances and Disease Registry (ATSDR)

I turn now to Dr. Ducatman's reliance on the ATSDR guidelines. Specifically, he relies upon the ATSDR criteria for medical monitoring of Flint, Michigan children. See his report (p. 21-3). He writes, using as support the ATSDR criteria, that:

> "The need for programs to provide diagnostic evaluation, intervention, amelioration, and treatment on a community wide basis is essential in Flint, and every exposed child should be eligible for preliminary assessment programs. Following such community-wide exposure, the follow-up need is to detect and, where detected and indicated to mitigate the undesirable outcomes of exposure."

As described in detail above, the documents Dr. Ducatman relies upon do not, in fact, provide evidence-based support for interventions to "mitigate" outcomes designed for and tested in populations of children exposed to lead.

It is important to point out that the ATSDR criteria involve tests for early detection and treatment. And the criteria state that early detection (i.e. screening) and the accompanying treatments or interventions involved must have been demonstrated to "interrupt the progress of disease, improve the quality of life, or be amenable to primary prevention."

It follows that Dr. Ducatman needs to show that the program he proposes for the Flint, Michigan children meets these ATSDR criteria. In other words, the Flint program needs to have specific early detection tests and specific interventions. In addition, these early detection tests and interventions must have critical evidence-based tests supporting their efficacy and effectiveness. Dr. Ducatman's reliance upon the PEHSU-AAP and CDC documents described above provide no such evidence. Finally, as discussed earlier, the U.S. Preventive Services Task Force (2019) does not recommend such a program for children exposed to lead.

In sum, the organizational documents Dr. Ducatman cites provide no valid and reliable support for the medical monitoring program he proposes.

Section 6.4.3 Dr. Ducatman's Misunderstanding of the Nature of Early Detection

In addition to Dr. Ducatman's failure to provide scientific and ethical support for the program he proposes, he appears to have a fundamental misunderstanding of the nature of early detection. In order to understand this issue, I turn to a recent legal case in which Dr. Ducatman cited the ATSDR's criteria for medical monitoring, the same criteria he relies upon in the Flint matter. Of particular interest is the fact that the ATSDR specifically states that medical monitoring is equivalent to the "early detection" of asymptomatic disease in individuals. As the ATSDR confirms, medical monitoring is the same as screening for disease.

> "This is consistent with the definition of screening as the presumptive identification of unrecognized disease or defect by the application of tests, examinations or other procedures which can be applied rapidly. Screening tests sort out apparently well persons who probably have a disease from those who probably do not. A screening test is not intended to be diagnostic. Persons with positive or suspicious findings must be referred to their physicians for diagnosis and necessary treatment."

With these criteria in mind, it will be important to examine how Dr. Ducatman applied these criteria to the need for a medical monitoring program in the case of a community allegedly exposed to PFOA, a chemical that allegedly contaminated drinking water. It is alleged in that same case, kidney and testicular cancer are alleged to have been caused by PFOA.

Dr. Ducatman states that early detection (i.e. screening) for these two cancers is beneficial. He writes (Report, p. 18) the following:

> "Early diagnosis and treatment of virtually all cancers is beneficial, and this is certainly true of kidney cancer and testicular cancer that are associated with PFOA exposures. (59-61, 128)."

This statement is scientifically and medically false. It is not true that early diagnosis and treatment is beneficial for "virtually all cancers." Documentation of this fact can be found at the National Cancer Institute's Physician Data Query (PDQ) website (www.cancer.gov/pdq). There it is clearly stated that there are cancers for which it has been established that "early diagnosis and treatment" is beneficial, for example, lung cancer (in smokers), cervical and breast cancer in women and colon cancer in men and women. However, the NCI-PDQ does not include kidney and testicular cancer in this regard.

The NCI PDQ also states that there is no standard or routine screening test for bladder cancer, thyroid cancer, endometrial cancer, liver cancer, and prostate cancer.

The American Cancer Society echoes the NCI-PDQ. At their website (www.cancer.org), the ACS states that none of the tests available for kidney cancer are recommended for screening.

Indeed, the only cancers that the American Cancer Society recommends for screening are: cervical, colorectal, and breast. They are circumspect about prostate and lung cancer.

> "The American Cancer Society's guidelines for average-risk adults recommend regular screening for breast cancer, cervical cancer, and colorectal cancer, based on scientific evidence that shows those screenings may help save lives."

What is most important here, however, is the fact that early detection and treatment for kidney cancer and testicular cancer has not been established, in direct contrast to what Dr. Ducatman opines. The American Cancer Society website states:

"There is no standard or routine screening test for testicular cancer."

Section 6.4.4 Summary of Dr. Ducatman's Report and Opinions

It follows that Dr. Ducatman's criteria for recommending screening tests and interventions are not evidence-based, in part because his views on early detection conflict with well-established organizational views on this topic, as described above. Although cancer is not an issue in this litigation, the basic principles of early detection apply to all exposure-disease relationships, including the lead-neurodevelopmental relationships. In the end, Dr. Ducatman is making uninformed claims about the nature of early detection—i.e. medical monitoring—programs.

The implication of this critical discussion of Dr. Ducatman's errors for the Flint Class Certification matter is clear. He appears to be unaware of the ways in which early detection tests and interventions are tested and recommended. None of the documents Dr. Ducatman relies upon recommend early detection tests and interventions for children who have been exposed to lead at levels likely experienced by members of the "Minors Subclass." It follows that Dr. Ducatman has no scientific basis for recommending screening and interventions for some if not most—perhaps all—members of the "Minors Subclass."

Finally, it is important to note that, despite his discussion of the PEHSU-AAP and CDC documents, Dr. Ducatman does not provide the details of testing and intervention(s) for the medical monitoring program he proposes for the "Minors Subclass."

Rather, he relies upon Dr. Keating "concerning approaches to childhood performance and behavioral testing, and appropriate intervention" (Ducatman Report, p. 9 and confirmed in Ducatman Deposition Vol. I, p. 129-30). For what Dr. Ducatman describes as "lead poisoning," Dr. Ducatman relies upon the expert reports of Drs. Hu and Lanphear (Ducatman Report, p. 9 and confirmed in Ducatman Deposition Vol. I, p. 129-30). It follows that if the opinions of Drs. Keating, Hu, and/or Lanphear are found to be invalid and unreliable, then Dr. Ducatman's opinions will be negatively affected. I will have more to say about the interconnections among these experts later in this report. For now, I turn to Dr. Keating's report and deposition. It is there that we, the readers, should expect to find a detailed description of the early detection tests to be included in the program proposed by Dr. Ducatman, how those tests have been tested to demonstrate their effectiveness as early detection tests, and what interventions have been implemented and found to be effective, i.e. more effective than testing done when symptoms arise. Recall that early detection programs are designed to identify individuals likely to have an outcome before they are symptomatic.

Finally, and most importantly, Dr. Ducatman provides no mention much less a discussion of the need for informed consent within which assurances that the program's risks outweigh its benefits would be communicated to the members of the "Minors Subclass" and their parents or guardians. As such, Dr. Ducatman's program fails the essential component of any early detection—medical monitoring program—based on well-established ethical principles.

Section 6.4.5 Summary of Dr. Ducatman's Report and Opinions

Dr. Ducatman's proposed program for medical monitoring lacks substance. He provides no evidence and no arguments to support the requirement that the program is based on the concept of "early detection," a concept required by the ATSDR guidelines he relies upon. Dr. Ducatman ignores the fact that organizational statements contradict his views. Dr. Ducatman relies upon Dr. Hu and Dr. Lanphear for the claim that the relationship between lead exposure and IQ as applied to the "Minors Subclass" is causal. But those two experts have failed to provide that support. Finally, Dr. Ducatman relies upon Dr.

Keating's opinions regarding the specific interventions to be included in the proposed medical monitoring program. As I will show in the next section, Dr. Keating fails to provide those interventions.

Finally, Dr. Ducatman provides no mention much less a full description of the need for informed consent wherein the risks and benefits of his medical monitoring program would be communicated to the individual members of the "Minors Subclass" and to their parents or guardians. He provides no information that the tests and interventions he proposes—or those presumably proposed by Dr. Keating—have been tested in adequately designed studies of children exposed to lead to show that, in fact, the benefits of the program outweigh the risks.

These errors of omission on the part of Dr. Ducatman are fatal. No medical monitoring program can be implemented without assurances in place that these ethical requirements are met.

For all these reasons and arguments above, I find Dr. Ducatman's opinions invalid and unreliable.

Section 6.5 Dr. Daniel P. Keating

Dr. Keating has submitted a 35-page report in this matter, dated June 29, 2020. He was deposed on October 29, 2020 and again on December 3, 2020. According to Dr. Keating, his purpose is to provide the following (Report, p. 2):

> "…expert opinions regarding the adverse neurodevelopmental and behavioral effects of lead poisoning, including the dangers to the health and function of children in Flint resulting from their exposure to lead in Flint's public drinking water. I also address the necessary remedies for these effects, including feasible assessment of neurodevelopmental harm and intervention efforts to mitigate them."

Inasmuch as Drs. Hu and Lanphear rather than Dr. Keating have as their primary responsibility opinions on the potential effects of lead on children, I will focus my attention on Dr. Keating's description of what he calls the "necessary remedies" for lead's effects and the "feasible assessment of neurodevelopmental harm," given that his opinions are to be incorporated into Dr. Ducatman's early detection—medical monitoring—program.

Simply put, Dr. Keating appears to have been engaged to provide the details for Dr. Ducatman's medical monitoring program, i.e. the early detection test(s) and the interventions Dr. Ducatman believes could mitigate the effects of lead allegedly caused by exposure to lead during the Flint Water Switch.

The purpose of my critique is to address the extent to which Dr. Keating provides evidence—i.e. reliable and valid scientific evidence—that the assessments and interventions he recommends for the Flint "Minors Subclass" are, in fact, consistent with the ATSDR guidelines on early detection methods. To be clear, it is important to examine whether these assessments and interventions have been demonstrated to be effective. Dr. Keating must show that the recommended assessments and interventions for asymptomatic children are better than the same assessments and interventions undertaken when the children have become symptomatic. As noted above, only randomized clinical trials and to a lesser extent, well-designed epidemiological studies, can demonstrate these conditions after being systematically reviewed.

Given that both the CDC document and the USPSTF document state that no such studies have been published, it seems difficult if not impossible for Dr. Keating to provide this essential information. Nevertheless, I will carefully examine what Dr. Keating has to say.

A major section of Dr. Keating's report is devoted to what he describes as an example of neurodevelopmental screening. In short, he describes a series of approaches to neurodevelopmental testing that purport to measure "domain of functioning," "executive function," "academic achievement," and "daily adaptive functioning." There is no mention much less a discussion of the extent to which these various approaches have been rigorously tested as early detection tests. Similarly, there is no mention of the actual tests to be administered.

In addition, when Dr. Keating describes the interventions to be used in this medical monitoring program—after the initial assessment—he states that "multiple clearinghouses (should be used) to identify 'what works,' to inform those decisions" (Keating Report, p. 27). In other words, Dr. Keating does not provide any specific interventions to, in his words, "mitigate" what he believes to be the effects of lead exposure on the "Minors Subclass." In the end, Dr. Keating provides no tests and no interventions to fill in the gaps of Dr. Ducatman's proposed medical monitoring program.

Summary: The Medical Monitoring Program Proposed by the Plaintiffs' Experts is Devoid of Critical Substance

In sum, there is no description much less discussion of how the medical monitoring program that is proposed by Dr. Ducatman will find early evidence of the effects of lead and no information at all on what specific interventions will be included. The program proposed by Drs. Ducatman and Keating is not a medical monitoring program at all. It has not been shown to be an effective early detection program as the ATSDR regulations require. In the end, it lacks so much substance that it cannot be considered scientifically or ethically appropriate.

Section 6.5.1 A Fundamental and Fatal Logical Flaw in Dr. Keating's Report and Not Addressed by Any Plaintiffs' Expert

Dr. Keating does not address the following fundamental problem: in his description, the assessment of the children will take place sometime in the future, say, 2021. Yet, how can the individuals administering the tests in 2021—i.e. tests which he does not specify—5 years after the Flint Water Switch know with any certainty that the results of the tests reveal effects of the Flint Water Switch? After all, if the members of the "Minors Subclass" were not given these tests—or similar versions—prior to the Flint Water Switch or immediately after the Flint Water Switch, then any adverse result found 5 years after the Flint Water Switch could have been caused by factors that occurred prior to the Flint Water Switch, during the Flint Water Switch or, alternatively, in the 5 years after the Flint Water Switch. My concern here sits at the very heart of good scientific practice and causal inference.

It appears that the plaintiffs—including but not limited to Dr. Keating—simply assume that if they observe a test result in 2021 for a member of the "Minors Subclass" that is deemed to be negative, e.g. an IQ of 95, then they know with certainty that the IQ of that same individual was higher before the Flint Water Switch and was unaffected by anything that occurred during the 5 years subsequent to the Flint Water Switch. These assumptions are so fraught with uncertainty that they cannot be accepted. In essence, the plaintiffs are attempting to argue that they can discern the cause of an effect by simply measuring an effect 5 years after the Flint Water Switch without taking into account alternative causes

before, during and after the Flint Water Switch. The problem only becomes more difficult given my discussion of the uncertainties and inaccuracies associated with the use of the Lanphear et al. (2005) pooled analysis which the plaintiffs incorrectly argue is, in essence, a "physical law" that precisely determines a decline in IQ for a given blood lead level.

To emphasize how unscientific and illogical the plaintiffs' arguments are, imagine the following scenario: assume that you take a medication for 3 months during 2014-2015 that is known to raise blood pressure. Assume that in 2021—approximately 5-6 years later—you take your blood pressure and discover that it is elevated. You do not know what your blood pressure was before 2014, during 2014-15, and for the 5-6 years up until 2021. How can you know—with any certainty—that the reason your blood pressure is elevated is due to (caused by) taking medication in 2014-2015? There is no way to know that. Any elevation in 2021 could have been due to measurement error, the medication taken in 2014-5, factors to which you were exposed during the 5-6 years after 2014-2015, including factors that reduce blood pressure (e.g. diet and exercise leading to weight loss or any number of other medications).

This scenario is precisely what the plaintiffs' experts would have us believe about the lead-neurodevelopmental (IQ) relationship. Without knowing anything about the members of the "Minors Subclass" prior to 2021 and assessing the neurodevelopment in 2021 the plaintiffs believe they can know that the findings were caused by the Flint Water Switch and not by anything else.

The plaintiffs' serious scientific problem does not get solved by claiming that the effects of lead are irreversible. Even if they are irreversible, then the effects might still have been caused by factors prior to, during, or after the Flint Water Switch, i.e. factors other than lead. Furthermore, if the effects truly are irreversible, then how can any intervention be helpful? The plaintiffs claim that their interventions—which are not specified—will "mitigate" the effects. But according to every dictionary I own, "mitigation" means to "reduce the severity or seriousness" of something. Irreversible effects cannot be mitigated if they are truly irreversible.

These logical errors provide one more reason to reject the plaintiffs' experts' claims.

Section 6.5.2 Summary of Dr. Keating's Opinions

Dr. Keating has failed to provide a key component of the medical monitoring program proposed by Dr. Ducatman. Dr. Keating has failed to provide any evidentiary support—much less a systematic assessment of evidence—for the claim that the neurodevelopment assessment he describes is an example of an early detection assessment.

Dr. Keating's opinions are invalid and unreliable.

Section 6.6 Summary of the Plaintiffs' Experts' Opinions: Returning to their Interdependencies

It has been shown that Dr. Lanphear's and Dr. Hu's opinions regarding the relationship between lead exposure and IQ are invalid and unreliable. Specifically, there is no good scientific rationale for claiming that every member of the "Minors Subclass" was harmed by the events of the Flint Water Switch, where that harm is defined in terms of a decline in IQ. There are too many uncontrolled factors that could explain declines in IQ and, as importantly, the plaintiffs do not know based on actual measurements which of the members of the "Minors Subclass" experienced a decline in IQ nor do they know how much

and to what extent each individual member of the "Minors Subclass" was exposed to lead before the Flint Water Switch, during the Flint Water Switch, and after the Flint Water Switch.

With regard to behavioral disorders, recall that Dr. Hu does not claim that these disorders (e.g. attention deficit hyperactivity disorder and its related symptoms) were caused by the hypothetical lead exposure potentially experienced by every member of the "Minors Subclass." As I described, he opines that each individual member of the "Minors Subclass" would need to undergo the process of differential diagnosis and etiologic assessment in order to discover which individual member experienced one of these outcomes. However, general causation relationships are required for Dr. Ducatman's medical monitoring program that is based upon the ATSDR guidelines. It follows that Dr. Ducatman's medical monitoring claims have failed to be supported by the opinions of Drs. Lanphear and Hu. Furthermore, Dr. Ducatman's medical monitoring program claims rely upon Dr. Keating's opinions on the specific evidence-based neurodevelopmental tests and interventions. However, Dr. Keating does not provide these. As described above, Dr. Keating has failed to show that the assessments he recommends have been adequately tested as early detection tests. Finally, Dr. Keating provides no specific interventions for Dr. Ducatman's medical monitoring program.

Without rehearsing the many other scientific errors found in these plaintiffs' experts' reports, it is clear that Dr. Ducatman's medical monitoring program is invalid and unreliable in part because the opinions of Drs. Lanphear, Hu, and Keating are invalid and unreliable.

PART SEVEN: OPINIONS

The plaintiffs' claim regarding class certification can be rejected on scientific grounds because they fail to take into consideration the fact that exposure to lead, the effects of lead, the effects of other causes of declines in IQ are unlikely, from a scientific perspective, to be identical or even similar across all members of the "Minors Subclass." Individual issues, in other words, outweigh commonalities. To be more specific, the failure of the plaintiffs' experts' case regarding class certification stems from the following:

1. Failure to account for the extent to which exposure to lead through drinking water—existence, duration, and concentration—and other sources of lead exposure occurred for individual members of the "Minors Subclass."
2. Failure to account for the possibility that factors other than lead exposure—the many environmental factors described above—were responsible for IQ deficits in the members of the "Minors Subclass" and that these factors are likely present in and affect each member of the subclass differently.
3. Failure to account for the relevance of individual measured blood lead levels of the members of the "Minors Subclass" and relying therefore solely on hypothetical data generated by an untested assumption-laden simulation modeling exercise. Note that blood lead levels are available for representative class plaintiffs; see Section 4.2 of this report.
4. Failure to resolve the relevant and important disagreement among two of the plaintiffs' experts, Drs. Lanphear and Hu on the issue of the shape of the relationship between blood lead levels and IQ decline.

For a discussion of these opinions regarding the plaintiffs' class certification claims, see PART ONE, Section 1.4.

Causal associations between exposure to low levels of blood lead (≤ 5 µg/dL) and neurodevelopmental outcomes (whether cognitive or behavioral) in children have not been established given the absence of effective adjustment for well-established causes and risk factors—i.e. confounders—of these neurodevelopmental outcomes in published epidemiological studies, the problems with the reliability of tests for low blood lead levels, and sparse epidemiological data. Note that this opinion negatively affects the viability and legitimacy of the plaintiffs' class certification claim. See also Section 4.2 of this report where blood lead levels of the representative class plaintiffs are tallied. All known values of blood lead levels in all representative class plaintiffs are below 5 µg/dL.

Causal associations between exposure to blood lead levels between 5 µg/dL and 10 µg/dL have not been established given the absence of effective adjustment for well-established causes and risk factors—i.e. confounders—of these neurodevelopmental outcomes in published epidemiological studies. This opinion also negatively affects the viability and legitimacy of the plaintiffs' class certification claim.

The pooled analysis by Lanphear et al. (2005) upon which the plaintiffs' experts rely is designed in such a way that does not provide a sufficient justification for establishing a causal relationship between blood lead levels and IQ in children. The design failures of this pooled analysis reflect the design problems of the original studies involved, namely, the use of concurrent blood lead and IQ measurements as well as failure to control for many established neurodevelopmental risk factors. As a result, the use of this analysis to predict IQ declines in individual members of the "Minors Subclass" is too uncertain to be scientifically valid and reliable.

The prevalence of causal factors and risk factors for adverse neurodevelopmental outcomes other than lead in women in Flint, Michigan during the years 2008-2015 is sufficient to explain the occurrence of those outcomes in children born during these years who were subsequently evaluated (e.g. in the years 2014 to the present). Put another way, given that the Lanphear et al. (2005) analysis fails to control for many of these factors—including but not limited to exposure to methylmercury, phthalates, PBDEs, and PCBs—it should not be used to make claims about the health of the children of Flint, Michigan who may or may not have been exposed to lead in water, depending upon their individual circumstances and who may have been exposed to a variety of other factors, i.e. well-established risk factors for neurodevelopmental outcomes other than lead.

Measured blood lead levels in young children (≤ 5 years of age) residing in Flint, Michigan during the switch from drinking water provided by the Detroit Water Authority (DWA) to the Flint River (FRW) revealed weak temporary elevations in average BLLs consistent with random variation. The average increase in BLLs among the Flint children was approximately 0.11 µg/dL, an increase that—if the relationship between BLLs and IQ is assumed to be causal—could be associated with changes in neurodevelopmental outcomes (e.g. IQ) so small as to be clinically insignificant and uninterpretable.

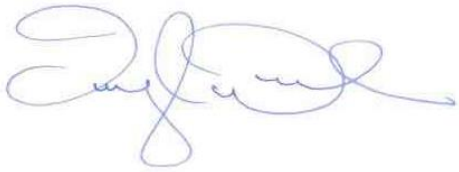
An individual's blood lead level is a result of exposure to all prior and current sources of lead in the environment over the lifespan of that individual. For children, sources of lead exposure include but are not limited to dust, soil, water, food, toys, and paint. The blood lead value measured in any child is also affected by the child's age, mouthing behaviors, socioeconomic situation (including the extent to which lead exists in the plumbing of the child's home), iron status, and ethnicity. The relationship between lead and IQ that the plaintiffs rely upon in this litigation is limited to that found in Lanphear et al. (2005) which is based on blood lead levels. Note the importance of individual rather than group (common) characteristics.

The opinions of Drs. Lanphear and Hu regarding general causation—i.e. opinions regarding the extent to which the relationship between exposure to lead and neurodevelopmental outcomes is causal—are scientifically invalid and unreliable on methodologic grounds. These will be described in more detail in Part Six of this report. In short, Drs. Lanphear and Hu do not provide systematic assessments of the scientific literature, fail to document claims with peer-reviewed scientific publications, and, in the case of Dr. Hu, make incorrect claims about the nature of risk.

The opinions of Drs. Ducatman and Keating regarding a medical monitoring program are scientifically invalid and unreliable. In addition, these experts have failed to provide any assurance that ethical imperatives of such programs—e.g. informed consent and the balance of benefits and risks of early detection—have been considered much less included. The medical monitoring program proposed by these experts cannot be said to be based on evidence that the benefits outweigh the risks. As such, it would be scientifically and ethically inappropriate to apply it to any individual member of the "Minors Subclass" much less the entire "Minors Subclass."

It has not been established that there is a causal relationship between exposure to low-levels of lead (e.g. < 10 µg/dL) and renal disease/renal dysfunction or between exposure to low-levels of lead and hypertension (increased blood pressure) in children. In addition, it has not been established that renal disease or hypertension diagnosed in adulthood can be causally linked to low-level exposure to lead in childhood. It has not been established that there is a causal relationship between exposure to low-levels of lead (e.g. < 10 µg/dL) as children and cardiovascular disease in adults. Similarly, it has not been established that there is a causal relationship between exposure to low-levels of lead as children and essential tremor in adults. See Appendices G, H, I, and J for details. It has not been established that low levels of lead (e.g. < 10 µg/dL) cause spontaneous abortions.

These opinions are made with a reasonable degree of scientific certainty.



January 5, 2021 Douglas L. Weed, M.D., M.P.H., Ph.D.

PART EIGHT: REFERENCES

Section 8.1 Methodological References

Agency for Toxic Substances and Disease Registry (ATSDR). ATSDR's Final Criteria for Determining the Appropriateness of a Medical Monitoring Program Under CERCLA. Federal Register Vol. 60, No. 145, July 28, 1995:388440-4.

Alexander DD, Weed DL, Cushing CA, Lowe KA. Meta-analysis of prospective studies of red meat consumption and colorectal cancer. Eur J Cancer Prevention 2011;20:293-307.

Alexander DD, Weed DL, Mink PJ, et al. A weight-of-evidence review of epidemiologic studies of colorectal cancer in pesticide applicators. Int Arch Occup Environ Health 2012;85:715-45.

Alexander DD, Weed DL, Chang ET, et al. A systematic review of multivitamin-multimineral use and cardiovascular disease and cancer incidence and total mortality. J Amer Coll Nutr 2013;32:339-54.

Alexander DD, Weed DL. On the need for improved methodological quality of published reviews. Am J Clin Nutr 2016;103:683-4.

Althuis MD, Weed DL, Frankenfeld CL. Evidence-based mapping of design heterogeneity prior to meta-analysis: a systematic review and evidence synthesis. Systematic Reviews 2014;3:80.

American College of Epidemiology Ethics Guidelines. Ann Epidemiol 2000;10:487-97.

Aschengrau A, Seage GR. The Approach and Evolution of Epidemiology. Chapter 1 in: Essentials of Epidemiology in Public Health. Sudbury, MA:Jones and Bartlett, 2003;1-32.

Aschengrau A, Seage GR. Cohort Studies. Chapter 8 in: Essentials of Epidemiology in Public Health. Sudbury, MA:Jones and Bartlett, 2003;193-220.

Aschengrau A, Seage GR. Case-Control Studies. Chapter 9 in: Essentials of Epidemiology in Public Health. Sudbury, MA:Jones and Bartlett, 2003;221-298.

Aschengrau A, Seage GR. Bias. Chapter 10 in: Essentials of Epidemiology in Public Health. Sudbury, MA:Jones and Bartlett, 2003;251-79.

Aschengrau A, Seage GR. Confounding. Chapter 11 in: Essentials of Epidemiology in Public Health. Sudbury, MA:Jones and Bartlett, 2003;281-98.

Aschengrau A, Seage GR. Random Error. Chapter 12 in: Essentials of Epidemiology in Public Health. Sudbury, MA:Jones and Bartlett, 2003;299-333.

Aschengrau A, Seage GR. Effect Measure Modification. Chapter 13 in: Essentials of Epidemiology in Public Health. Sudbury, MA:Jones and Bartlett, 2003;334-47.

Aschengrau A, Seage GR. The Epidemiologic Approach to Causation. Chapter 15 in: Essentials of Epidemiology in Public Health. Sudbury, MA:Jones and Bartlett, 2003;375-401.

Baker D, Nieuwenhuijsen MJ. Environmental epidemiology: Study methods and application. New York:Oxford University Press, 2009.

Beaglehole R, Bonita R, Kjellstrom T. Causation in Epidemiology. Chapter 5 in: Basic Epidemiology. Geneva:World Health Organization, 1993;71-81.

Bhandari M, Devereqeux PJ, Montori V, et al. Users' guide to the surgical literature: how to use a systematic literature review and meta-analysis. J Can Chir 2004;47:60-7.

Bhopal R. What is epidemiology? Chapter 1 in: Concepts of Epidemiology. New York:Oxford University Press, 2002;1-16.

Bhopal R. Cause and effect: The epidemiological approach. Chapter 5 in: Concepts of Epidemiology. New York:Oxford University Press, 2002;98-132.

Bind MA. Causal modeling in environmental health. Annu Rev Public Health 2019;40:23-43.

Boffetta P, Farioli A, Rizzello E. Application of epidemiological findings to individuals. Med Lav 2020;111:10-21.

Bramwell VHC, Williams CJ. Do authors of review articles use systematic methods to identify, assess, and synthesize information? Ann Oncol 1997;8:1185-95.

Breslow RA, Ross SA, Weed DL. Quality of reviews in epidemiology. Am J Pub Health 1998;88(3):475-7.

Campbell MJ, Machin D. Medical Statistics: A Commonsense Approach. Chichester:John Wiley & Sons. 1990.

Charlton BG. Screening, ethics, and the law. BMJ 1992;305:521.

Checkoway H, Pearce N, Kriebel D. Research Methods in Occupational Epidemiology. Oxford: University Press, 2004.

Cogliano VJ, Baan RA, Straif K, et al. The science and practice of carcinogen identification and evaluation. Environ Health Perspect 2004;112:1269-74.

Cogliano VJ, Baan RA, Straif K, et al. Use of mechanistic data in IARC evaluations. Environ Molec Mutag 2008;49:100-9.

Cole P. Causality in epidemiology, health policy, and law. Environmental Law Reporter 1997;27:10279-85.

Crowther MA, Cook DJ. Trials and tribulations of systematic reviews and meta-analyses. Hematology 2007;493-7.

Dominici F, Zigler C. Best practices for gauging evidence of causality in air pollution epidemiology. Amer J Epidemiol 2017;186:1303-9.

Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomized and non-randomized studies of health care interventions. J Epidemiol Commun Health 1998;52:377-84.

Duncan M, Moschopoulou E, Herrington E, et al. Review of systematic reviews of non-pharmacological interventions to improve quality of life in cancer survivors. BMJ Open 2017;Nov 28;7(11):e015860.

Edwards PJ, Hall DMB. Screening, ethics, and the law. BMJ 1992;305;267-8.

Evans AS. Causation and disease: a chronological journey: The Thomas Parran Lecture. Am J Epidemiol 1978;108:148-58.

Golder S, Loke Y, McIntosh HM. Poor reporting and inadequate searches were apparent in systematic reviews of adverse effects. J Clin Epidemiol 2008;61:440-8.

Goldman SA. Limitations and strengths of spontaneous reports data. Clin Therapeut 1998;20(Suppl C):C40-C44.

Goodman SN, Samet JM. Causation and Causal Inference. Chapter 1 in: Schottenfeld D, Fraumeni Jr, JF. Cancer Epidemiology and Prevention, 3rd ed. New York:Oxford, 2006;3-9.

Goodstein D. How science works. Reference Manual on Scientific Evidence. 3rd ed. Washington DC:National Academies Press, 2011:37-54.

Gordis L. From Association to Causation: Deriving Inferences from Epidemiologic Studies. Chapter 13 in: Epidemiology, 2nd ed. Philadelphia:W.B. Saunders, 2000;184-203.

Greenhalgh T. How to read a paper: Papers that summarise other papers (systematic reviews and meta-analyses). BMJ 1997;315:672-5.

Guirguis-Blake J, Calonge N, Miller T, et al. Current processes of the U.S. Preventive Services Task Force refining evidence-based recommendation development. Ann Intern Med 2007;147:117-22.

Guzelian PS, Guzelian CP. Authority based explanation. Science 2004;303:1468-9.

Hasan H, Muhammed T, Yu J, et al. Assessing the methodological quality of systematic reviews in radiation oncology: a systematic review. Ca Epidemiol 2017;50:141-9.

Hill AB. The Environment and Disease: Association or Causation? Proc Roy Soc Med 1965;58:295-300.

Hill AB. Statistical evidence and inference. Chapter 24 in: A Short Textbook of Medical Statistics. London:Hodder and Stoughton 1971:283-96.

Holland PW. Statistics and causal inference. J Am Stat Assoc 1986;81:945-60.

Holland WW. Screening: reasons to be cautious. BMJ 1993;306:1222-3.

Hutchison BG. Critical appraisal of review articles. Can Fam Physician 1993;39:1097-102.

Institute of Medicine. Knowing what works in health care: a roadmap for the nation. January, 2008. At: www.nap.edu.

Jick H. The discovery of drug-induced illness. NEJM 1977;296:481-5.

Kelsey JL, Pettiti DB, King AC. Key Methodologic Concepts and Issues. Chapter 2 in: Brownson RR, Pettiti DB. Applied Epidemiology: Theory to Practice. New York:Oxford University Press, 1998;35-69.

Kleinbaum DG, Kupper LL, Morgenstern H. Fundamentals of Epidemiologic Research. Chapter 2 in: Epidemiologic Research. Belmont, CA:Lifetime Learning. 1982;19-39.

Koopman JS, Weed DL. Epigenesis theory: A mathematical model relating causal concepts of pathogenesis in individuals to disease patterns in populations. Am J Epidemiol. 1990;132:366-90.

Kriebel D, Tickner J. Reenergizing public health through precaution. Am J Public Health 2001;91:1351-5.

Krimsky S. The Weight of Evidence in Policy and Law. Am J Pub Health 2005;95(Suppl 1):S129-S136.

Kung J, Chiappelli F, Cajulis OO, et al. From systematic reviews to clinical recommendations for evidence-based health care: validation of revised assessment of multiple systematic reviews (R-AMSTAR) for grading of clinical relevance. Open Dent J 2010;4:84-91.

Li L, Ying XJ, Sun TT, et al. Overview of methodological quality of systematic reviews about gastric cancer risk and protective factors. Asian Pac J Cancer Prev 2012;13:20169-79.

Liberati A, Altman DG, Tegzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. Ann Int Med 2009;151:65-94.

Lichtenstein AH, Yetley EA, Lau J. Application of systematic review methodology to the field of nutrition. J Nutrition 2008;138:2297-306.

Lundh A, Knijnenburg SL, Jorgensen AW, et al. Quality of systematic reviews in pediatric oncology—a systematic review. Cancer Treat Rev 2009;35:645-52.

MacMahon B, Pugh TF. Concepts of Cause. Chapter 2 in: Epidemiology: Principles and Methods. Boston:Little, Brown, 1970.

Manchikanti L, Datta S, Smith HS et al. Evidence-based medicine, systematic reviews, and guidelines in interventional pain management: Part 6. Systematic reviews and meta-analyses of observational studies. Pain Physician 2009;12:819-50.

Mant D, Fowler G. Mass screening: theory and ethics. BMJ 1990;300:916-8.

Marshall KG. Prevention. How much harm? How much benefit? 4. The ethics of informed consent for preventive screening programs. Can Med Assoc J 1996;155:377-83.

Mausner JD, Bahn AK. The Search for Causal Relations: Observational Studies. Chapter 5 in: Epidemiology: An Introductory Text. Philadelphia:W.B. Saunders. 1974;91-111.

McDonald JH. Handbook of Biological Statistics, 3rd ed. Sparky House Publishing, Baltimore, Maryland, 2014.

McKeown RE, Weed DL. Ethics in epidemiology and public health. II. Applied terms. J Epidemiol Commun Health 2002;56:739-41.

McLaren L, Hawe P. Ecological perspectives in health research. J Epidemiol Commun Health 2005;59:6-14.

Mignini LE, Khan KS. Methodological quality of systematic reviews of animal studies: a survey of reviews of basic research. BMC Med Research Methodology 2006, 6:10. doi:10.1186/1471-2288-6-10.

Milne R, Chambers L. Assessing the scientific quality of review articles. J Epidemiol Community Health 1993;47(3):169-70.

Moher D, Tsertsvadze A, Tricco AC, et al. When and how to update systematic reviews. Cochrane Database Syst Rev 2008;23(1):MR000023.

Moher D, Tetzlaff J, Tricco AC, et al. Epidemiology and reporting characteristics of systematic reviews. PLoS Med 2007;4(3): e78. doi:10.1371/journal.pmed.0040078.

Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Medicine 2009;6:e1000097.

Montori VM, Swiontkowski MF, Cook DJ, et al. Methodologic issues in systematic reviews and meta-analyses. Clin Orthop Rel Rsch 2003;413:43-54.

Morris MC, Evans DA, Hebert LE, et al. Methodological issues in the study of cognitive decline. Amer J Epidemiol 1999;149:789-93.

Muir Gray JA. Two classes of creativity—improving systematic reviews. J Epidemiol Commun Health 1994;48:4-5.

Mullen PD, Ramirez G. The promise and pitfalls of systematic reviews. Annu Rev Public Health 2006;27:81-102.

Mulrow CD. The medical review article: state of the science. Ann Intern Med 1987;106:485-8.

Mulrow CD. Rationale for systematic reviews. BMJ 1994;309:597-9.

National Research Council (NRC). Toxicity Testing in the 21st Century: A Vision and a

Strategy. Committee on Toxicity Testing and Assessment of Environmental Agents, National Research Council. Washington DC,National Academies Press, 2007.

Noordzij M, Hooft L, Dekker FW et al. Systematic reviews and meta-analyses: when they are useful and when to be careful. Kidney Int 2009;76:1130-6.

Oxman AD, Guyatt GH. Guidelines for reading literature reviews. Can Med Assoc J 1988;138:697-703.

Oxman DA. Checklists for review articles. Brit Med J 1994;309:648-51.

Oxman AD, Schunemann HJ, Fretheim A. Improving the use of research evidence in guideline development: 8. Synthesis and presentation of evidence. Health Research Policy and Systems 2006, 4:20. doi:10.1186/1478-4505-4-20.

Parascandola M, Weed DL. Causation in epidemiology. J Epidemiol Commun Health 2001;55:905-12.

Paul KC, Haan M, Mayeda ER, et al. Ambient air pollution, noise, and late-life cognitive decline and dementia risk. Annu Rev Public Health 2019;40:203-20.

Petticrew M. Systematic reviews from astronomy to zoology: myths and misconceptions. Brit Med J 2001;322:98-101.

Porta M. Dictionary of Epidemiology. 5th ed. New York:Oxford University Press, 2008.

Reference Manual on Scientific Evidence. 2nd ed. Federal Judicial Center. 2000.

Reference Manual on Scientific Evidence. 3rd ed. Federal Judicial Center and the National Research Council of the National Academies of Science. National Academies Press, Washington DC. 2011.

Rochon PA, Bero LA, Bay AM, et al. Comparison of review articles published in peer-reviewed and throwaway journals. JAMA 2002;287:2853-6.

Rosenstock 2005. Textbook of Clinical Occupational and Environmental Medicine. 2nd ed.

Rothman KJ. Causal Inference in Epidemiology. Chapter 2 in: Modern Epidemiology. Boston: Little, Brown. 1986;7-21.

Rothman KJ. What is Causation? Chapter 2 in: Epidemiology: an Introduction. New York:Oxford University Press, 2002;8-23.

Rothman KJ and Greenland S. Causation and Causal Inference. Chapter 2 in: Modern Epidemiology, 2nd ed. Philadelphia:Lippincott, Raven. 1998.

Sachs RM, Bornichak EA. An evaluation of spontaneous adverse drug reaction monitoring systems. Am J Med 1986;81:49-55.

Sackett DL. Screening in family practice: prevention, levels of evidence and the pitfalls of common sense. J Fam Pract 1987;24:233-4.

Sanderson S, Tatt ID, Higgins JPT. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. Int J Epidemiol 2007;36:666-76.

Schreider J, Barrow C, Birchfield N, et al. Enhancing the credibility of decisions based on scientific conclusions: transparency is imperative. Tox Sci 2010;doi: 10.1093/toxsci/kfq102.

Schwartz PH, Meslin EM. The ethics of information: absolute risk reduction and patient understanding of screening. J Gen Intern Med 2008;23:867-70.

Shea BJ, Bouter LM, Peterson J, et al. External validation of a measurement tool to assess systematic reviews (AMSTAR). PLoS One 2007;2(12):e1350.

Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Research Methodology 2007, 7:10. doi:10.1186/1471-2288-7-10.

Shea BJ, Hamel C, Wells GA, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. J Clin Epidemiol 2009;62:1013-20.

Stoate HG. Can health screening damage your health? J Roy Coll Gen Pract 1989;39:193-5.

Straus SE, McAlister FA. Evidence-based medicine: a commentary on common criticisms. CMAJ 2000;163:837-41.

Stroup DF, Berlin JA, Morton SC. Meta-analysis of observational studies in epidemiology: a proposal for reporting. J Am Med Assoc 2000;283:2008-12

Susser M. Rules of inference in epidemiology. Reg Tox Pharm 1986;6:116-28.

Surgeon General's Advisory Committee on Smoking and Health. Smoking and health: 1964. Rockville MD: U.S. Public Health Service, 1964 (DHEW Publication no. (PHS) 1103).

United States Environmental Protection Agency (EPA). Guidelines for Carcinogen Risk Assessment. Risk Assessment Forum U.S. Environmental Protection Agency, Washington, DC. EPA/630/P-03/001F March, 2005.

United States Preventive Services Task Force (USPSTF). Procedure Manual. Section 4: Evidence Report Development. AHRQ Publication #08-05118-EF, July, 2008.

Vandenbroucke JP. Case reports in an evidence-based world. J Roy Soc Med 1999;92:159-63.

Venning GR. Identification of adverse reactions to new drugs. III. Alerting processes and early warning systems. BMJ 1983;286:458-60.

Vetter N, Matthews I. Causation. Chapter 3 in: Epidemiology and Public Health Medicine. London: Churchill, Livingstone, 1999;23-30.

Volmink J, Siegfried N, Robertson K, et al. Research synthesis and dissemination as a bridge to knowledge management: the Cochrane Collaboration. Bull WHO 2004;82:778-83.

Weed DL. On the logic of causal inference. Am J Epidemiol 1986;123:965-79.

Weed DL, Tyroler, HA, Shy CM. The healthy worker effect in actively-working communications workers. J Occup Med 1987;29:335-9.

Weed DL. Causal and Preventive Inference. Chapter 17 in: Greenwald P, Kramer BS, Weed DL. Cancer Prevention and Control. New York:Marcel Dekker, 1995;285-302.

Weed DL, Gorelic LS. The practice of causal inference in cancer epidemiology. Cancer Epidemiol Biomark Prev 1996;5:303-11.

Weed DL. Methodological guidelines for review papers. J Natl Cancer Inst 1997;89:6-7.

Weed DL. Preventing scientific misconduct. Am J Public Health 1998;88:125-9.

Weed DL, Hursting SD. Biologic plausibility in causal inference: current method and practice. Am J Epidemiol 1998;147:415-25.

Weed DL. Ethics and consent. Chapter 6 in: Kramer BS, Gohagan JK, Prorok PC. Cancer Screening: Theory and Practice. New York:Marcel Dekker, 1999:89-140.

Weed DL. Epidemiological evidence and causal inference. Hemat/Oncol Clin N Amer 2000;14:797-807.

Weed DL. Interpreting epidemiological evidence: how meta-analysis and causal inference methods are related. Int J Epidemiol 2000;29:387-90.

Weed DL. Methods in epidemiology and public health: does practice match theory? J Epidemiol Commun Health 2001;55:104-10.

Weed DL, McKeown RE. Ethics in epidemiology and public health. I. Technical terms. J Epidemiol Community Health 2001;55:855-7.

Weed DL. Environmental epidemiology: basics and proof of cause-effect. Toxicology 2002; 181-182:399-403.

Weed DL. Weight of evidence: review of concept and methods. Risk Anal 2005;25:1545-57.

Weed DL. Evidence synthesis and general causation: key methods and an assessment of reliability. Drake Law Review 2006;54:639-650.

Weed DL. The nature and necessity of scientific judgment. J Law Policy 2007;15:135-64.

Weed DL. Conflicts of interest. J Epidemiol Commun Health 2009;63:601-2.

Weed DL. Meta-analysis and causal inference: a case study of benzene and non-Hodgkin's lymphoma. Ann Epidemiol 2010;20:347-55.

Weed DL, Althuis MA, Mink PJ. Quality of reviews on sugar-sweetened beverages and health outcomes. Am J Clin Nutr 2011;94:1340-7.

Weed DL. The quality of nutrition and cancer reviews: a systematic assessment. Crit Rev Food Sci Nutrition 2013;53:276-86.

Weed DL. Causal inference in epidemiology: potential outcomes, pluralism, and peer review. Int J Epidemiol 2016;45:1838-40.

Weed DL. Analogy in causal inference: rethinking Austin Bradford Hill's neglected consideration. Ann Epidemiol 2018;28:343-346.

Weed DL. The need for systematic reviews in oncology. JNCI J Natl Cancer Inst 2018;110(8): djy050.

Weir E, Schabas R, Wilson K, et al. A Canadian framework for applying the precautionary principle in public health issues. Can J Public Health 2010;101:396-8.

Wells GA, Shea B, O'Connelle D, et al. The Ottawa Hospital Research Institute. Our Research. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomized studies in meta-analyses. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp [accessed March, 2020].

Wynder EL, Higgins IT, Harris RE. The wish bias. J Clin Epidemiol 1990;43:619-21.

Section 8.2 Additional References

Section 8.2.1 References Cited Regarding Background Information on Lead and Neurodevelopmental Outcomes (Cognitive and Behavioral)

Agency for Toxic Substances and Disease Registry (ATSDR). Toxicological Profile for Lead. Draft for Public Comment. May, 2019. Final (August, 2020) Centers for Disease Control and Prevention.

Allotey J, Zamora J, Cheong-See F, et al. Cognitive, motor, behavioral and academic performances of children born preterm: a meta-analysis and systematic review involving 64 061 children. Brit J Obs Gyn 2018;125:16-25.

Azeredo A, Moreira D, Barbosa D. ADHD, CD, and ODD: Systematic review of genetic and environmental risk factors. Research Developmental Disabilities 2018:82;10–19.

Bishop EG, Cherny SS, Corley R, et al. Development genetic analysis of general cognitive ability from 1 to 12 years in a sample of adoptees, biological siblings, and twins. Intelligence 2003;31:31-49.

Braaten EB, Norman D. Intelligence (IQ) testing. Pediatrics Rev 2008;27:403-7.

Braun JM, Kahn RS, Froehlich T, et al. Exposures to environmental toxicants and attention deficit hyperactivity disorder in US children. Environ Health Perspect 2006;114:1904-9.

Breslau N, Chilcoat HD, Susser ES, et al. Stability and change in children's intelligence quotient scores: A comparison of two socioeconomically disparate communities. Am J Epidemiol 2001;154:711-7.

Buettner C, Mukamal KJ, Gardiner P, et al. Herbal supplement use and blood lead levels of United States adults. J Gen Intern Med 2009;24:1175-82.

Caldwell KL, Cheng PY, Jarrett JM, et al. Laboratory measurement implications of decreasing childhood blood lead levels. Pediatrics 2017;140:doi:10.1542/peds.2017-0272.

Centers for Disease Control and Prevention (CDC). Fourth Report on Environmental Chemicals. 2009.

Cortese S, Tessari L. Attention-Deficit/Hyperactivity Disorder (ADHD) and Obesity: Update 2016. Curr Psychiatry Rep 2017;19:4 DOI 10.1007/s11920-017-0754-1.

Crump KS, Van Landingham C, Bowers TS, et al. A statistical reevaluation of the data used in the Lanphear et al. (2005) pooled-analysis that related low levels of blood lead to intellectual deficits in children. Crit Rev Toxicol 2013;43:785-99.

Del-Ponte B, Quinte GC, Cruz S, et al. Dietary patterns and attention deficit/hyperactivity disorder (ADHD): A systematic review and meta-analysis. J Affective Disorders 2019;252:160-73.

Dong T, Hua W, Zhou X, et al. Prenatal exposure to maternal smoking during pregnancy and attention-deficit/hyperactivity disorder in offspring: A meta-analysis. Reprod Toxicol 2017;76:63-70.

Donzelli G, Llopis-Gonzalez A, Llopis-Morales A, et al. Particulate Matter Exposure and Attention-Deficit/Hyperactivity Disorder in Children: A Systematic Review of Epidemiological Studies. Int J Environ Research Pub Health 2020;17:67.

Dooling EC. Cognitive disorders in children. Curr Opin Pediatr 1993;5:675-9.

Drover SM, Villanger GD, Aase H, et al. Maternal Thyroid Function During Pregnancy or Neonatal Thyroid Function and Attention Deficit Hyperactivity Disorder A Systematic Review. Epidemiology 2019;30:130-44.

Edwards M, Triantafyllidou S, Best D. 2009. Elevated blood lead in young children due to lead-contaminated drinking water: Washington, DC, 2001-2004. Environ Sci Technol 2009;43: 1618-23.

Fetene DM, Betts KS, Alati R. Maternal thyroid dysfunction during pregnancy and behavioural and psychiatric disorders of children: a systematic review. Eur J Endocrinol 2017;77:R261-73.

Flannery BM, Dolan LC, Hoffman-Pennesi D, et al. U.S. Food and Drug Administration's interim reference levels for dietary lead exposure in children and women of childbearing age. Reg Tox Pharm 2020;110:104516.

Fletcher JM, Stuebing KK, Hughes LC. The IQ scores should be corrected for the Flynn effect in high stakes decisions. J Psychoeduc Assessment 2010;28:469-73.

Flynn JR. The mean IQ of Americans: Massive gains 1932-1978. Psych Bull 1984;95:29-51.

Franz AP, Bolat GU, Bolat H, et al. Attention-Deficit/Hyperactivity Disorder and Very Preterm/Very Low Birth Weight: A Meta-analysis. Pediatrics 2018;141(1):e20171645.

Fuertes E, Standl M, Forns J, et al. Traffic-related air pollution and hyperactivity/inattention, dyslexia and dyscalculia in adolescents of the German GINIplus and LISAplus. Environ International 2016;97:85-92.

Gottfried AW, Gottfried AE, Guerin DW. The Fullerton Longitudinal Study: A long-term investigation of intelligence and motivational giftedness. J Educ Gifted 2006;29:430-50.

Gou X, Wang Y, Tang Y, et al. Association of maternal prenatal acetaminophen use with the risk of attention deficit/hyperactivity disorder in offspring: A meta-analysis. Aust NZ J Psychiatry 2019;53:195–206.

Haier RJ. Increased intelligence is a myth (so far). Front Systems Neurosci 2014;8:1-3.

Hu H. "Heavy Metal Poisoning." Chapter 449 in: Jameson JL, Fauci AS, Kasper DL, et al. eds. 20th Edition Harrison's Principles of Internal Medicine. New York:McGraw Hill 2018;449:3297-3300.

Huang L, Wang Y, Zhang L, et al. Maternal Smoking and AttentionDeficit/Hyperactivity Disorder in Offspring: A Meta-analysis. Pediatrics 2018;141(1):e20172465.

Johnson NH, Ash KO, Nuttall KL, et al. The adequacy of capillary specimens for determining whole blood lead. Ann Clin Lab Sci 1997;27:179-84.

Jones RL, Homa DM, Meyer PA, et al. Trends in blood lead levels and blood lead testing among US children aged 1 to 5 years, 1988-2004. Pediatrics 2009;123:e376-e385.

Khoshbakht Y, Bidaki R, Salehi-Abargouei, et al. Vitamin D Status and Attention Deficit Hyperactivity Disorder: A Systematic Review and Meta-Analysis of Observational Studies. Adv Nutr 2018;9:9-20.

Kim S, Arora M, Fernandez C, et al. Lead, mercury, and cadmium exposure and attention deficit hyperactivity disorder in children. Environ Rsch 2013;126:105-10.

Lam J, Lanphear BP, Bellinger D, et al. Developmental PBDE Exposure and IQ/ADHD in Childhood: A Systematic Review and Meta-analysis. Environ Health Perspect 2017; 086001.

Lanphear BP, Matte TD, Rogers J, et al. The contribution of lead-contaminated house dust and residential soil to children's blood lead levels. Environ Research Section A 1998;79:51-68.

Lanphear BP, Hornung R, Ho M, et al. Environmental lead exposure during early childhood. J Pediatr 2002;140:40-7.

Lanphear BP, Hornung R, Khoury J, et al. Low-level environmental lead exposure and children's intellectual function: An international pooled analysis. Environ Health Perspect 2005;113:894-9.

Mahaffey KR, Annest JL, Roberts J, et al. National estimates of blood lead levels: United States, 1976-1980. Association with selected demographic and socioeconomic factors. New Engl J Med 1982;307:573-9.

Maher GM, O'Keeffe GW, Kearney PM, et al. Association of Hypertensive Disorders of Pregnancy With Risk of Neurodevelopmental Disorders in Offspring A Systematic Review and Meta-analysis. JAMA Psychiatry 2018;75:809-19.

McCall RB et al. Transitions in infant sensorimotor development and the prediction of childhood IQ. Amer Psychologist 1972;27:728-48.

McClure LF, Niles JK, Kaufman HW. Blood lead levels in young children: US, 2009-2015. J Pediatrics 2016;175:173-81.

Morbidity and Mortality Weekly Report (MMWR). Blood lead levels-United States, 1988-1991. 1994;43:545-8.

National Toxicology Program (NTP). NTP Monograph on Health Effects of Low-Level Lead. 2012 U.S. Department of Health and Human Services, Office of Health Assessment and Translation, Division of the National Toxicology Program, National Institute of Environmental Health Sciences, National Institutes of Health, U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES.

Neisser U, Boodoo G, Bouchard TJ Jr. et al. Intelligence: knowns and unknowns. Amer Psychologist 1996;51:77-101.

Nielsen TM, Pedersen MV, Milidou I, et al. Long-term cognition and behavior in children born at early term gestation: A systematic review. Acta Obstet Gynecol Scand 2019;98:1227-34.

Nigg JT, Knottnerus M, Martel MM, et al. Low blood lead levels associated with clinically diagnosed attention-deficit/hyperactivity disorder and mediated by weak cognitive control. Biol Psychiatry 2008;63:325-31.

Parsons PKJ, Reilly AA, Esernio-Jenssen D. Screening children exposed to lead: an assessment of the capillary blood lead fingerstick test. Clin Chem 1997;43:302-11.

Pelham WE Jr., Fabiano GA, Massetti GM. Evidence-based assessment of attention deficit hyperactivity disorder in children and adolescents. J Clin Child Adolesc Psychology 2005;34:449-76.

Ramsden S, Richardson FM, Josse G, et al. Verbal and non-verbal intelligence changes in the teenage brain. Nature 2011;479:113-6.

Schneider W, Niklas F, Schmiedeler S. Intellectual development from early childhood to early adulthood: The impact of early IQ differences on stability and change over time. Learning Individual Differences 2014;32:156-62.

Scott JG, Mihalopoulos C, Erskine HE, et al. Childhood mental and developmental disorders. Chapter 8 in: Mental, Neurological, and Substance Use Disorders. Patel V, Chisholm D, Dua T, et al. (eds.) Disease Control Priorities, 3rd ed. Washington DC:World Bank, 2015:145-61.

Sentis A, Sunyer J, Dalmau-Buenoa A, et al. Prenatal and postnatal exposure to NO2 and child attentional function at 4–5 years of age. Environ Int 2017;106:170-7.

Shoaff JR, Coull B, Weuve J, et al. Association of Exposure to Endocrine-Disrupting Chemicals During Adolescence With Attention-Deficit/Hyperactivity Disorder–Related Behaviors. JAMA Net Open 2020;3(8):e2015041.

Taylor L, Jones RL, Ashley K, et al. Comparison of capillary earlobe and venous blood monitoring for occupational lead surveillance. J Lab Clin Med 2004;143:217-24.

Thygesen M, Holsta GJ, Hansene B, et al. Exposure to air pollution in early childhood and the association with Attention-Deficit Hyperactivity Disorder. Environ Research 2020;183:108930.

Trahan L, Stueving KK, Hiscock MK, et al. The Flynn Effect: A meta-analysis. Psychol Bull 2014;140:1332-60.

Tsoi MF, Cheung CL, Cheung TT, et al. Continual decrease in blood lead level in Americans: United States National Health Nutrition and Examination Survey, 1999-2014. Am J Med 2016;129:1213-8.

Van Dongen J, Zilhao NR, Sugden K, et al. Epigenome-wide Association Study of AttentionDeficit/Hyperactivity Disorder Symptoms in Adults. Biol Psychiatry 2019;86:599-607.

Verner MA, Plusquellec P, Desjardins JL, et al. Prenatal and early-life polychlorinated biphenyl (PCB) levels and behavior in Inuit preschoolers. Environ Int 2015;78:90-4.

Wetherill L, Foroud T, Goodlett C. Meta-Analyses of Externalizing Disorders: Genetics or Prenatal Alcohol Exposure? Alcohold Clin Exp Research 2018;42:162-72.

Zhang T, Sidorchuk A, Sevilla-Cermeno et al. Association of Cesarean Delivery With Risk of Neurodevelopmental and Psychiatric Disorders in the Offspring A Systematic Review and Meta-analysis. JAMA Netw Open 2019;2(8):e1910236.

Section 8.2.2 References Specific to Flint, Michigan and the Flint Water Switch

Note: The references in this section 8.2.2 are included for completeness but those relevant to the measurement of lead exposure in Flint residents are primarily cited in the text of the report.

Abokifa AA, Katz L, Sela L. Spatiotemporal trends of recovery from lead contamination in Flint, MI as revealed by crowdsourcing water sampling. Water Res 2019;171:115442.

Abouk R, Adams S. Birth outcomes in Flint in the early stages of the water crisis. J Public Health Pol 2018;39:68-85.

Abuelaish I, Russel KK. The Flint water contamination crisis: the corrosion of positive peace and human decency. Medicine Conflict Survival 2017;33:242-9.

Allen JM, Cuthbertson AA, Liberatore HK, et al. Showering in Flint, MI: Is there a DPB problem? J Environ Sci 2017;58:271-84.

Banner W. "Toxicohistrionics": Flint, Michigan and the lead crisis. J Pediatrics 2018;197:15-6.

Baum R, Bartram J, Hrudey S. The Flint water crisis confirms that U.S. drinking water needs improved risk management. Environ Sci Technol 2016;50:5436-7.

Belllinger DC. Lead contamination in Flint—An abject failure to protect public health. New Engl J Med 2016;374:1101-3.

Campbell C, Greenberg R, Mankikar D, et al. A case study of environmental injustice: The failure in Flint. Environ Res Pub Health 2016;13:1-11.

Carravallah LA, Reynolds LA, Woolford SJ. Lessons for physicians from Flint's water crisis. AMA J Ethics 2017;19:1001-10.

Craft-Blackshear MG. Lessons learned from the crisis in Flint, Michigan regarding the effects of contaminated water on maternal and child health. JOGNN 2017;46:258-66.

Cuthbertson CA, Newkirk C, Ilardo J, et al. Angry, scared, and unsure: Mental health consequences of contaminated water in Flint, Michigan. J Urban Health 2016;93:899-908.

DeWitt RD. Pediatric lead exposure and the water crisis in Flint, Michigan. J Am Acad Physician Asst 2017;30:43-6.

Fortenberry GZ, Reynolds P, Burrer SL, et al. Assessment of behavioral health concerns in the community affected by the Flint water crisis—Michigan (USA) 2016. Prehosp Disaster Med 2018;33:256-65.

Garner E, Brown CL, Schwake DO, et al. Comparison of whole-genome sequences of Legionella pneumophila in tap water and in clinical strains, Flint, Michigan, USA, 2016. Emerg Infect Dis 2019;25:2013-20.

Gomez HF, Borgialli DA, Sharman M, et al. Blood lead levels of children in Flint, Michigan: 2006-2016. J Pediatrics 2018;197:158-64.

Gomez HF, Borgialli DA, Sharman M, et al. Blood lead levels in females of childbearing age in Flint, Michigan, and the water crisis. Obstet Gynecol 2019;134:628-35.

Gomez HF, Borgialli DA, Sharman M, et al. Analysis of blood lead levels of young children in Flint, Michigan before and during the 18-month switch to Flint River water. Clin Toxicol 2019;57:790-7.

Goovaerts P. The drinking water contamination crisis in Flint: Modeling trends of lead level since returning to Detroit water system. Sci Total Environ 2017;581-2:66-79.

Goovaerts P. Monitoring the aftermath of Flint drinking water contamination crisis: Another case of sampling bias? Sci Total Environ 2017;590-591:139-53.

Gostin LO. Politics and public health: The Flint drinking water crisis. Hasting Center Rpt 2016;July-August:5-6.

Greenberg MR. Delivering fresh water: critical infrastructure, environmental justice, and Flint, Michigan. AJPH 2016;106:1358-60.

Grossman DW, Slusky DJG. The effect of an increase in lead in the water system on fertility and birth outcomes: the case of Flint, Michigan. 2017.

Grossman DW, Slusky DJG. The impact of the Flint Water Crisis on fertility. Demography 2019;56:2005-31.

Hanna-Attisha M, LaChance J, Sadler RC, et al. Elevated blood levels in children associated with the Flint drinking water crisis: a spatial analysis of risk and public health response. Am J Pub Health 2016;106:283-90.

Hanna-Attisha M. Flint kids: tragic resilient, and exemplary. Am J Public Health 2017;107:651-2.

Heard-Garris NJ, Roche J, Carter P, et al. Voices from Flint: community perceptions of the Flint water crisis. J Urban Health 2017;94:776-9.

Jennings B, Duncan LL. Water safety and lead regulation: physicians' community health responsibilities. AMA J Ethics 2017;19:1027-35.

Kennedy C, Yard E, Dignam T, et al. Blood lead levels among children aged <6 years—Flint, Michigan, 2013-2016. MMWR 2016;65:650-4.

Key KD. Expanding ethics review processes to include community-level protections: a case study from Flint, Michigan. AMA J Ethics 2017;19:989-98.

Kruger DJ, Kodjebacheva GD, Cupal S. Poor tap quality experiences and poor sleep quality during the Flint, Michigan municipal water crisis. Sleep Health 2017;3:241-3.

Laidlaw MAS, Filippelli GM, Sadler RC, et al. Children's blood lead seasonality in Flint, Michigan (USA) and soil-sourced lead hazard risks. Int J Environ Res Public Health 2016;13:358.

Lieberman A. Recovery efforts in Flint slowly begin to take form. Lancet 2016;387:1499-1500.

Liu G, Zhang Y, Knibbe WJ, et al. Potential impacts of changing supply-water quality on drinking water distribution: a review. Water Rsch 2017;116:135-48.

Markel H. Remember Flint. Milbank Quarterly 2016;94:229-36.

Masten SJ, Davies SH, McElmurry SP. Flint water crisis: what happened and why? J Am Water Works Assoc 2016;108:22-34.

Mayfield KE, Carolan M, Weatherspoon L, et al. African American women's perception on access to food and water in Flint, Michigan. J Nutr Educ Behav 2017;49:519-25.

Miller A, Yeskey K, Garantziotis S, et al. Integrating health research into disaster response: the new NIH disaster research response program. Int J Environ Res Public Health 2016;13:676.

Nelson R. Crisis in Flint: lead and Legionnaires' disease. Lancet 2016;16:298-9.

Newland J. The lead crisis in Flint, Michigan. Nurse Practitioner 2016;41:12.

Pieper KJ, Tang M, Edwards MA. Flint water crisis caused by interrupted corrosion control: investigating "Ground Zero" home. Environ Sci Technol 2017;51:2007014.

Pieper KJ, Martin R, Tang M, et al. Evaluating water lead levels during the Flint water crisis. Environ Sci Technol 2018;52:8124-32.

Reynolds K. Images of health and learning. Waterborne. AMA J Ethics. 2017;19:1036-42.

Rosner D. Flint, Michigan: A century of environmental injustice. Am J Pub Health 2016;106:200-1.

Rosner D. A lead poisoning crisis enters its second century. Health Affairs 2016;35:5.

Ruckart PZ, Ettinger AS, Hanna-Attisha M, et al. The Flint water crisis: a coordinated public health emergency response and recovery initiative. JPHMP 2019;25:S84-90.

Sadler RC, LaChance J, Hanna-Attisha M. Social and built environmental correlates of predicted blood lead levels in the Flint water crisis. Am J Pub Health 2017;107:763-9.

Sadler RC, Furr-Holden D. The epidemiology of opioid overdose in Flint and Genesee County, Michigan: Implications for public health practice and intervention. Drug Alcohol Depend 2019;204:107560.

Shah KK, Oleske JM, Gomez HF, et al. Blood lead concentrations of children in the United States: a comparison of states using two large databases. J Pediatrics 2017;185:218-23.

Silbergeld EK. Drinking water and the developing brain. Cerebrum July, 2016:1-15.

Smith AF, Huss A, Dorevitch S, et al. Multiple sources of the outbreak of Legionnaires' disease in Genesee County, Michigan, in 2014 and 2015. Environ Health Perspect 2019;127(12)127001-1 to 11.

Zahran S, Mielke HW, McElmurry SP, et al. Determining the relative importance of soil sample locations to predict risk of child lead exposure. Environ Int 2013;60:7-14.

Zahran S, McElmurry SP, Sadler RC. Four phases of the Flint water crisis: evidence from blood lead levels in children. Environ Research 2017;157:160-72.

Zahran S, McElmurry SP, Kilgore PE, et al. Assessment of the Legionnaires' disease outbreak in Flint, Michigan. PNAS 2018;e1730-9.

Section 8.2.3 References for the "Other Contaminants" Issue

Causality

Engel SM, Wolff MS. Causal inference considerations for endocrine disruptor research in children's health. Annu Rev Public Health 2013;34:139-58.

Confounding and other Methodologic Issues

Bellinger DC. Assessing environmental neurotoxicant exposures and child neurobehavior. Epidemiol 2004;15:383-4.

Bellinger DC. A strategy for comparing the contributions of environmental chemicals and other risk factors to neurodevelopment of children. Environ Health Perspect 2012;120:501-7.

Bradley RH, Caldwell BM, Rock SL, et al. Home Observation for Measurement of the Environment: Development of a home inventory for use with families having children 6 to 10 years old. Contemp Educ Psychol 1988;13:58-71.

Lanphear BP, Hornung RW, Khoury J, et al. The conundrum of unmeasured confounding: Comment on: "Can some of the detrimental neurodevelopmental effects attributed to lead be due to pesticides? By Brian Gulson." Sci Total Environ 2008;396:196-200.

Mink PJ, Goodman M, Barraj LM, et al. Evaluation of uncontrolled confounding in studies of environmental exposures and neurobehavioral testing in children. Epidemiol 2004;15:385-93.

Studies (Exposure: Smoking and/or Alcohol)

Evlampidou I, Bagkeris M, Vardavas C, et al. Prenatal second-hand smoke exposure measured with urine cotinine may reduce gross motor development at 18 months of age. J Pediatrics 2015;167:246-52.

Han JY, Kwon HJ, Ha M, et al. The effects of prenatal exposure to alcohol and environmental tobacco smoke on risk for ADHD: A large population-based study. Psychiatry Res 2015;225:164-8.

Kesmodel US, Kjaersgaard MIS, Denny CH, et al. The association of pre-pregnancy alcohol drinking with child neuropsychological functioning. BJOG 2014;doi:10.1111/1471-0528.

Moore BF, Shapiro AL, Wilkening G, et al. Prenatal exposure to tobacco and offspring neurocognitive development in the Healthy Start Study. J Pediatrics 2019;1-7.

Polanska K, Muszynski P, Sobala W, et al. Maternal lifestyle during pregnancy and child psychomotor development—Polish Mother and Child Cohort Study. Early Hum Develop 2015;91:317-25.

Polanska K, Krol A, Merecz-Kot D, et al. Environmental tobacco smoke exposure during pregnancy and child neurodevelopment. Int J Environ Res Pub Health 2017;14:796.

Wiebe SA, Clark CAC, De Jong DM, et al. Prenatal tobacco exposure and self-regulation in early childhood: Implications for developmental psychotherapy. Develop Psychopathol 2015;27:397-409.

Studies (Exposure: Weight—Overweight/Obesity)

Krzeczkowski JE, Boylan K, Arbuckle TE, et al. Neurodevelopment in 3-4 year old children exposed to maternal hyperglycemia or adiposity in utero. Early Hum Develop 2018;125:8-16.

Nichols AR, Rundle AG, Factor-Litvak P, et al. Prepregnancy obesity is associated with lower psychomotor development scores in boys at age 3 in low-income, minority birth cohort. J Develop Orig Health Dis 2020;11:49-57.

Widen EM, Nichols AR, Kahn LG, et al. Prepregnancy overweight and obesity are associated with impaired child neurodevelopment. Matern Child Nutr 2018;14:doi:10.1111/mcn.12481.

Widen EM, Nichols AR, Kahn LG, et al. Prepregnancy obesity is associated with cognitive outcomes in boys in a low-income, multiethnic birth cohort. BMC Pediatr 2019;19:507.

Yeung EH, Sundaram R, Ghassabian A, et al. Parental obesity and early childhood development. Pediatrics 2017;139:e20161459.

Studies (Exposure: Pesticides)

Cartier C, Warembourg C, Maner-Idrissi GL, et al. Organophosphate insecticide metabolites in prenatal and childhood urine samples and intelligence scores at 6 years of age: results from the Mother-Child PELAGIE cohort (France). Environ Health Perspect 2016;124:674-80.

Coker E, Gunier R, Bradman A, et al. Association between pesticide profiles used on agricultural fields near maternal residences during pregnancy and IQ at age y years. Int J Environ Res Pub Health 2017;14:506; doi:10.3390/ijerph14050506.

Donauer S, Altaye M, Xu Y, et al. An observational study to evaluate associations between low-level gestational exposure to organophosphate pesticides and cognition during early childhood. Am J Epidemiol 2016;184:410-8.

Engel SM, Bradman A, Wolff MS, et al. Prenatal organophosphorus pesticide exposure and child neurodevelopment at 24 months: An analysis of four birth cohorts. Environ Health Perspect 2016;124:822-30.

Fluegge KR, Nishioka M, Wilkins III JR. Effects of simultaneous prenatal exposures to organophosphate and synthetic pyrethroid insecticides on infant neurodevelopment at three months of age. J Environ Toxicol Public Health 2016;1:60-73.

Furlong MA, Herring A, Buckley JP, et al. Prenatal exposure to organophosphorus pesticides and childhood neurodevelopmental phenotypes. Environ Res 2017;158:737-47.

Furlong MA, Barr DB, Wolff MS, et al. Prenatal exposure to pyrethroid pesticides and childhood behavior and executive functioning. Neurotoxicol 2017;62:231-8.

Gunier RB, Bradman A, Harley KG, et al. Prenatal residential proximity to agricultural pesticide use and IQ in 7-year old children. Environ Health Perspect 2017;057002.

Quiros-Alcala L, Mehta S, Eskenazi B. Pyrethroid pesticide exposure and parental report of learning disability and attention deficit/hyperactivity disorder in U.S. children: NHANES 1999-2002. Environ Health Perspect 2014;122:1336-42.

Viel JF, Warenbourg C, Maner-Idrissi GL, et al. Pyrethroid insecticide exposure and cognitive developmental disabilities in children: the PELAGIE mother-child cohort. Environ Int 2015;82:69-75.

Viel JF, Rouget F, Warembourg C, et al. Behavioral disorders in 6-year-old children and pyrethroid insecticide exposure: the PELAGIE mother-child cohort. Occup Environ Med 2017;74:175-81.

Zhang Y, Han S, Liang D, et al. Prenatal exposure to organophosphate pesticides and neurobehavioral development of neonates: a birth cohort study in Shenyang, China. PLoS One 2014;9:2:e88491.

Studies (Exposure: Phthalates)

Dong R, Wu Y, Chen J, et al. Lactational exposure to phthalates impaired the neurodevelopmental function of infants at 9 months in a pilot prospective study. Chemosphere 2019;226:351-9.

Goodman JM, Ingle ME, Domino SE, et al. First trimester maternal exposures to endocrine disrupting chemicals and metals and fetal size in the Michigan Mother Infant Pairs study. J Dev Orig Health Dis 2019;10:447-58.

Hyland C, Mora AM, Kogut K, et al. Prenatal exposure to phthalates and neurodevelopment in the CHAMACOS cohort. Environ Health Perspect 2019;127:107010-1.

Jankowska A, Polanska K, Koch HM, et al. Phthalate exposure and neurodevelopmental outcomes in early school age children from Poland. Environmental Research 2019;179:108829.

Jankowska A, Polanska K, Hanke W, et al. Prenatal and early postnatal phthalate exposure and child neurodevelopment at age of 7 years—Polish Mother and Child Cohort. Environmental Research 2019;177:108626.

Jones B, Han TL, Delplancke T, et al. Association between maternal exposure to phthalates and lower language ability in offspring derived from hair metabolome analysis. Sci Rep 2018;8:6745.

Qian X, Li J, Xu S, et al. Prenatal exposure to phthalates and neurocognitive development in children at two years of age. Environment Int 2019;131:105023.

Radke EG, Braun JM, Nachman RM, et al. Phthalate exposure and neurodevelopment: A systematic review and meta-analysis of human epidemiological evidence. Environ Int 2020;137:105408.

Reyes JM, Price PS. Temporal trends in exposures to six phthalates from biomonitoring data: Implications for cumulative risk. Environ Sci Technol 2018;52:12475-83.

Tanner EM, Hallerback MU, Wikstrom S, et al. Early prenatal exposure to suspected endocrine disruptor mixtures is associated with lower IQ at age seven. Environment Int 2020;134:105185.

Varshavsky JR, Morell-Frosch R, Woodruff TJ, et al. Dietary sources of cumulative phthalates exposure among the U.S. general population in NHANES 2005-2014. Environ Int 2018;115:417-29.

Watkins DJ, Milewski S, Domino SE, et al. Maternal phthalate exposure during early pregnancy and at delivery in relation to gestational age and size at birth: A preliminary analysis. Reprod Toxicol 2016;65:59-66.

Zhang Q, Chen XZ, Huang X, et al. The association between prenatal exposure to phthalates and cognition and neurobehavior of children-evidence from birth cohorts. Neurotoxicol 2019;73:199-212.

Zota AR, Calafat AM, Woodruff TJ. Temporal trends in phthalate exposures: Findings from the National Health Examination Survey, 2001-2010. Environ Health Perspect 2014;122:235-41.

Studies (Exposure: Polybrominated Diphenyl Ethers—PBDE)

Batterman SA, Chernyak S, Jia C, et al. Concentrations and emissions of polybrominated diphenyl ethers from U.S. houses and garages. Environ Sci Technol 2009;43:2693-700.

Batterman SA, Godwin C, Chernyak S, et al. Brominated flame retardants in offices in Michigan, USA. Environ Int 2010;36:548-56.

Bradley PW, Wan Y, Jones PD, et al. PBDEs and methoxylated analogues in sediment cores from two Michigan, USA, inland lakes. Environ Tox Chem 2011;30:1236-42.

Braun JM, Yolton K, Stacy SL, et al. Prenatal environmental chemical exposures and longitudinal patterns of child neurobehavior. Neurotoxicology 2017;62:192-9.

Chen A, Yolton K, Rauch SA, et al. Prenatal polybrominated diphenyl ether exposures and neurodevelopment in U.S. children through 5 years of age: the HOME study. Environ Health Perspect 2014;122:856-62.

De Water E, Curtin P, Zilverstand A, et al. A preliminary study on prenatal polybrominated diphenyl ether serum concentrations and intrinsic functional network organization and executive functioning in childhood. J Child Psychol Psychiatry 2019;60:1010-20.

Ji H, Liang H, Wang Z, et al. Associations of prenatal exposures to low levels of polybrominated diphenyl ether (PBDE) with thyroid hormones in cord plasma and neurobehavioral development in children at 2 and 4 years. Environ Int 2019;131:105010.

Lenters V, Iszatt N, Forns J, et al. Early-life exposure to persistent organic pollutants (OCPs, PBDEs, PCBs, PFASs) and attention-deficit/hyperactivity disorder: a multi-pollutant analysis of a Norwegian birth cohort. Environ Int 2019;125:33-42.

Liang H, Vuong AM, Xie C, et al. Childhood polybrominated diphenyl ether (PBDE) serum concentration and reading ability at 5 and 8 years: The HOME Study. Environ Int 2019;122:330-9.

Orta OR, Wesselink AK, Bethea TN, et al. Correlates of plasma concentrations of brominated flame retardants in a cohort of U.S. black women residing in Detroit, Michigan. Sci Tot Environ 2020;714:136777.

Rice CP, Chernyak SM, Begnoche L, et al. Comparisons of PBDE composition and concentration in fish collected from the Detroit River, MI, and Des Plains River, IL. Chemosphere 49;2002:731-7.

Song W, Li A, Ford JC, et al. Polybrominated diphenyl ethers in the sediments of the Great Lakes. 2. Lakes Michigan and Huron. Environ Sci Technol 2005;39:3474-9.

U.S. EPA. Technical Fact Sheet—Polybrominated diphenyl ethers (PBDEs), November 2017.

Venier M, Dove A, ROmanak K, et al. Flame retardants and legacy chemicals in Great Lakes' water. Environ Sci Technol 2014;48:9563-72.

Vuong AM, Yolton K, Webster GM, et al. Prenatal polybrominated diphenyl ether and perfluoroalkyl substance exposures and executive function in school-age children. Environ Res 2016;147:556-64.

Vuong AM, Braun JM, Yolton K, et al. Prenatal and postnatal polybrominated diphenyl ether exposures and visual spatial abilities in children. Environ Res 2017;153:83-92.

Vuong AM, Yolton K, Poston KL, et al. Childhood polybrominated diphenyl ether (PBDE) exposure and executive function in children in the HOME study. Int J Hyg Environ Health 2018;221:87-94.

Yuan Y, Meeker JD, Ferguson KK. Serum polybrominated diphenyl ether (PBDE) concentrations in relation to biomarkers of oxidative stress and inflammation: The National Health and Nutrition Examination Survey 2003-2004. Sci Tot Environ 2017;575:400-5.

Zhang H, Yolton K, Webster GM, et al. Prenatal PBDE and PCB exposure and reading, cognition, and externalizing behavior in children. Environ Health Perspect 2017;125:746-52.

Zhu LY, Hites RA. Temporal trends and spatial distributions of brominated flame retardants in archived fish from the Great Lakes. Environ Sci Technol 2004;38:2779-84.

Zhu LY, Hites RA. Brominated flame retardants in sediment cores from Lakes Michigan and Erie. Environ Sci Technol 2005;39:3488-94.

Studies (Exposure: Polychlorinated Biphenyls—PCBs)

Caspersen IH, Aase H, Biele G, et al. The influence of maternal dietary exposure to dioxins and PCBs during pregnancy on ADHD symptoms and cognitive functions in Norwegian preschool children. Environ Int 2016;94:649-60.

Goodrich JM, Ingle ME, Domino SE, et al. First trimester maternal exposures to endocrine disrupting chemicals and metals and fetal size in the Michigan Mother Infant Pairs study. J Dev Orig Health Dis 2019;10:447-58.

Han L, Hsu WW, Todem D, et al. In utero exposure to polychlorinated biphenyls is associated with decreased fecundity in daughters of Michigan female fisheaters: a cohort study. Environ Health 2016;15:92.

Kim S, Eom S, Kim HJ, et al. Association between maternal exposure to major phthalates, heavy metals, and persistent organic pollutants, and the neurodevelopmental performances of their children at 1 to 2 years of age—CHECK cohort study. Sci Tot Environ 2018;624:377-84.

Kyriklaki A, Vafeiadi M, Kampouri M, et al. Prenatal exposure to persistent organic pollutants in association with offspring neuropsychological development at 4 years of age: The Rhea mother-child cohort, Crete, Greece. Environ Int 2016;97:204-11.

Lenters V, Iszatt N, Forns J, et al. Early-life exposure to persistent organic pollutants (OCPs, PBDEs, PCBs, PFASs) and attention-deficit/hyperactivity disorder: A multi-pollutant analysis of a Norwegian birth cohort. Environ Int 2019;125:33-42.

Lynch CD, Jackson LW, Kostyniak PJ, et al. The effect of prenatal and postnatal exposure to polychlorinated biphenyls and child neurodevelopment at age twenty four months. Reproductive Toxicol 2012;34:451-6.

Nakajima S, Saijo Y, Miyashita C, et al. Sex-specific differences in effect of prenatal exposure to dioxin-like compounds on neurodevelopment in Japanese children: Sapporo chort study. Environ Res 2017;159:222-31.

Rasmussen PW, Schrank C, Williams MCW. Trends of PCB concentrations in Lake Michigan coho and chinook salmon, 1975–2010. J Great Lakes Rsch 2014;40:748-54.

Ruel MVM, Bos AF, Soechitram SD, et al. Prenatal exposure to organohalogen compounds and children's mental and motor development at 19 and 30 months of age. Neurotoxicology 2019;72:6-14.

Salamova A, Pagano JJ, Holsen TM, et al. Post-1990 temporal trends of PCBs and organochlorine pesticides in the atmosphere and in fish from Lakes Erie, Michigan, and Superior. Environ Sci Technol 2013;47:9109-114.

Tsai MS, Chen MH, Lin CC, et al. Children's environmental health based on birth cohort studies of Asia. Sci Tot Environ 2017;609:396-408.

Wang BL, Pang ST, Sun JP, et al. Levels of polychlorinated biphenyls in settled house dust from urban dwellings in China and their neurodevelopmental effects of preschool-aged children. Sci Total Environ 2015;505:402-8.

Studies (Exposure: Fluoride)

Saeed M, Rehman MYA, Farooqi A, et al. Co-exposure effects of arsenic and fluoride on intelligence and oxidative stress in school-aged children: a cohort study. Environ Research 2020;110168.

Till C, Green R, Flora D, et al. Fluoride exposure from infant formula and child IQ in a Canadian birth cohort. Environ Int 2020;134:105315.

Wang M, Liu L, Li H, et al. Thyroid function, intelligence and low-moderate fluoride exposure among Chinese school-age children. Environ Int 2020;134:105229.

Reviews, Systematic Reviews, and Meta-Analyses (Other Contaminants)

Allotey J, Zamora J, Cheong-See F, et al. Cognitive, motor, behavioral and academic performances of children born preterm: a meta-analysis and systematic review involving 64,061 children. Brit J Obstet Gynecol 2018;125:16-23.

Alvarez-Bueno C, Cavero-Redondo I, Lucas-de la Cruz L, et al. Association between pre-pregnancy overweight and obesity and children's neurocognitive development: a systematic review and meta-analysis of observational studies. Int J Epidemiol 2017;1653-66.

Ben-Shlomo Y, Kuh D. A life course approach to chronic disease epidemiology: conceptual models, empirical challenges, and interdisciplinary perspectives. Int J Epidemiol 2002;31:285-93.

Burnett AC, Cheong JLY, Doyle LW. Biological and social influences on the neurodevelopmental outcomes of preterm infants. Clin Perinatol 2018;45:485-500.

Burns CJ, McIntosh LJ, Mink PJ, et al. Pesticide exposure and neurodevelopmental outcomes: review of the epidemiologic and animal studies. J Toxicol Environ Health, Pt B. 2013;16:127-283.

Chen R, Clifford A, Lang L, et al. Is exposure to secondhand smoke associated with cognitive parameters of children and adolescents?—a systematic literature review. Ann Epidemiol 2013;23:652-61.

Clifford A, Lang L, Chen R. Effects of maternal cigarette smoking during pregnancy on cognitive parameters of children and young adults: a literature review. Neurotox Teratol 2012;34:560-70.

Clifford A, Lang L, Chen R, et al. Exposure to air pollution and cognitive functioning across the life course—a systematic literature review. Environ Research 2016;147:383-98.

Duan Q, Jiao J, Chen X, et al. Association between water fluoride and the level of children's intelligence: a dose-response meta-analysis. Public Health 2018;154:87-97.

Ejaredar M, Nyanza EC, Eycke KT, et al. Phthalate exposure and childrens neurodevelopment: a systematic review. Environ Research 2015;142:51-60.

Fox J. A Life Course Approach to Chronic Disease Epidemiology (Book Review). J Epidemiol Commun Health 1998;317:421.

Gibson EA, Siegel EL, Eniola F, et al. Effects of polybrominated diphenyl ethers on child cognitive, behavioral, and motor development. Int J Environ Research Pub Health 2018;15:1636/doi:10.3390/ijerph15081636.

Gonzalez-Alzaga B, Lacasana M, Aguilar-Garduno C, et al. A systematic review of neurodevelopmental effects of prentatal and postnatal organophosphate pesticide exposure. Toxicol Lett 2014;230:104-21.

Grandjean P, Landrigan PJ. Developmental neurotoxicity of industrial chemicals. Lancet 2006;368:2167-78.

Grandjean P, Landrigan PJ. Neurobehavioral effects of developmental toxicity. Lancet Neurol 2014;13:330-8.

Grandjean P, Herz KT. Trace elements as paradigms of developmental neurotoxicants: Lead, methylmercury, and arsenic. J Trace Elements Med Biol 2015;31:130-4.

Grandjean P, Kishi R, Kogevinas M, et al. Prevention of developmental neurotoxicity. Epidemiology 2017;28:157-8.

Guth S, Huser S, Roth A, et al. Toxicity of fluoride: critical evaluation of evidence for human developmental neurotoxicity in epidemiological studies, animal experiments, and in vitro analyses. Arch Toxicol 2020;94:1375-1415.

Herting MM, Younan D, Campbell CE. Outdoor air pollution and brain structure and function from across childhood to young adulthood: A methodological review of brain MRI studies. Frontiers Pub Health 2019;17:1-21.

Hertz-Piccioto I, Sass JB, Engel S, et al. Organophosphate exposures during pregnancy and child neurodevelopment: recommendations for essential policy reforms. PLoS One 2018;15:e1002671.

Jurewicz J, Hanke W. Exposure to phthalates: reproductive outcome and children health: a review of epidemiologic studies. Int J Occup Med Environ Health 2011;24:115-41.

Jurewicz J, Polanska K, Hanke W. Exposure to widespread environmental toxicants and children's cognitive development and behavioral problems. Int J Occup Med Environ Health 2013;26:185-204.

Koureas M, Tsakalof A, Tsatsakis A, et al. Systematic review of biomonitoring studies to determine the association between exposure to organophosphorus and pyrethroid insecticides and human health outcomes. Toxicol Lett 2012;210:155-68.

Kuh D, Ben-Shlomo Y, Lynch J, et al. Life course epidemiology. J Epidemiol Commun Health 2003;57:778-83.

Lam J, Lanphear BP, Bellinger D, et al. Developmental PBDE exposure and IQ/ADHD in childhood: a systematic review and meta-analysis. Environ Health Perspect 2017;e086001.

Lanphear BP. The impact of toxins on the developing brain. Ann Rev Public Health 2015;36:211-30.

Lee DW, Kim MS, Lim YH, et al. Prenatal and postnatal exposure to di-(2-ethylhexyl) phthalate and neurodevelopmental outcomes: A systematic review and meta-analysis. Environ Research 2018;167:558-66.

Levy RJ. Carbon monoxide pollution and neurodevelopment: a public health concern. Neurotox Teratol 2015;49:31-40.

Linsell L, Malouf R, Morris J, et al. Prognostic factors for poor cognitive development in children born very preterm or with very low birth weight—a Systematic Review. JAMA Pediatr 2015;169:1162-72.

Mikolajewska K, Stragierowica J, Gromadzinska J. Bisphenol A—application, sources of exposure and potential risks in infants, children and pregnant women. Int J Occup Med Environ Health 2-15;28:209-41.

Munoz-Quezada MT, Lucero BA, Barr DB, et al. Neurodevelopmental effects in children associated with exposure to organophosphate pesticides: A systematic review. Neurotoxicol 2013;39:158-68.

Polanska K, Jurewicz J, Hanke W. Review of current evidence on the impact of pesticides, polychlorinated biphenyls and selected metals on attention deficit/hyperactivity disorder in children. IJOMEH 2013;26:16-38.

Polanska K, Jurewicz J, Hanke W. Smoking and alcohol drinking during pregnancy as the risk factors for poor child neurodevelopment-a review of epidemiological studies. Int J Occup Med Environ Health 2015;28:419-43.

Radke EG, Braun JM, Nachman RM, et al. Phthalate exposure and neurodevelopment: A systematic review and meta-analysis of human epidemiological evidence. Environ Int 2020;137:105408.

Rice D, Barone SJ. Critical periods of vulnerability for the developing nervous system: evidence from humans and animal models. Environ Health Perspect 2000;103(Suppl 5):511-33.

Sanchez CE, Barry C, Sabhlok A, et al. maternal pre-pregnancy obesity and child neurodevelopmental outcomes: a meta-analysis. Obesity Rev 2018;19:464-84.

Schug TT, Blawas AM, Gray K, et al. Elucidating the links between endocrine disruptors and neurodevelopment. Endocrinology 2015;156:1941-51.

Silbergeld EK. Drinking water and the developing brain. Cerebrum July 2016.

Smith LM, Santos LS. Review: Prenatal exposure: The effects of prenatal cocaine and methamphetamine exposure on the developing child. Birth Defects Rsch (Part C) 2016;108:142-6.

Spittle B. Reviews of developmental fluoride neurotoxicity by Grandjean and Guth et al. Fluoride 2020;53:204-19.

Sunyer J, Dadvand P. Pre-natal brain development as a target for urban air pollution. Basic Clin Pharmacol Toxicol 2019;125(Suppl 3):81-8.

Tsai MS, Chen MH, Lin CC, et al. Children's environmental health based on birth cohort studies of Asia (2)—air pollution, pesticides, and heavy metals. Environ Research 2019;179:108754.

Turyk ME, Bhavsar SP, Bowerman W, et al. Risks and benefits of consumption of Great Lakes fish. Environ Health Perspect 2012;120:11-8.

Twilhaar ES, de Kieviet JF, Aarnoudse-Moens CSH, et al. Academic performance of children born preterm: a meta-analysis and meta-regression. Arch Dis Child Fetal Neonatal Ed 2018;103:F322-F330.

Van der Burg JW, Sen S, Chomitz VR, et al. The role of systemic inflammation linking maternal body mass index to neurodevelopment in children. Pediatr Res 2016;79:3-12.

Vrijheid M, Casas M, Gascon M, et al. Environmental pollutants and child health—a review of recent concerns. Int J Hygiene Environ Health 2016;219:331-42.

Vuong AM, Yolton K, Dietrich KN, et al. Exposure to polybrominated diphenyl ethers (PBDEs) and child behavior: Current findings and future directions. Hormones Behavior 2018;101:94-104.

Weiss B. Vulnerability of children and the developing brain to neurotoxic hazards. Environ Health Perspect 2000;108(Suppl 3):375-81.

Wolff MS, Buckley J, Engel SM, et al. Emerging exposures of developmental toxicants. Curr Opin Pediatr 2017;29:218-24.

Ylijoki MK, Ekholm E, Ekblad M, et al. Prenatal risk factors for adverse developmental outcome in preterm infants—systematic review. Front Psychol 2019;10:595.

Zhang Q, Chen XZ, Huang X, et al. The association between prenatal exposure to phthalates and cognition and neurobehavior of children—evidence from birth cohorts. Neurotoxicol 2019;73:199-212.

Section 8.2.4 References regarding Exposure to Lead and Cognitive/Behavioral Outcomes

Al-Saleh I, Shinwari N, Nester M, et al. Longitudinal study of prenatal and postnatal lead exposure and early cognitive development in Al-Kharj, Saudi Arabia: a preliminary results of cord blood lead levels. J Tropical Pediatrics 2008;54:300-7.

Arbuckle TE, Davis K, Boylan K, et al. Bisphenol A, phthalates and lead and learning and behavioral problems in Canadian children 6-11 years of age: CHMS 2007-2009. Neurotoxicol 2016;54:89-98.

Baghurst PA, McMichael AJ, Wigg NR, et al. Environmental exposure to lead and children's intelligence at the age of seven years: the Port Pirie Cohort Study. New Engl J Med 1992;327:1279-84.

Bellinger DC, Stiles KM, Needleman HL. 1992. Low-level lead exposure, intelligence and academic achievement: a long-term follow-up study. Pediatrics 1992;90:855-61.

Bellinger D, Leviton A, Allred E, et al. Pre- and postnatal lead exposure and behavior problems in school-aged children. Environmental Research 1994;66:12-30.

Bellinger DC, Needleman HL. Intellectual impairment and blood lead levels. New Engl J Med 2003;349:500–2.

Blackowicz MJ, Hryhorczuk DO, Rankin KM, et al. The impact of low-level lead toxicity on school performance among Hispanic subgroups in the Chicago public schools. International Journal of Environmental Research Public Health 2016;13:774-85.

Braun JM, Kahn RS, Froehlich T, et al. Exposures to environmental toxicants and attention deficit hyperactivity disorder in U.S. children. Environ Health Perspect 2006;114:1904-9.

Braun J, Hoffman E, Schwartz J, et al. 2012. Assessing windows of susceptibility to lead-induced cognitive deficits in Mexican children. Neurotoxicology 2012;33:1040-7.

Canfield RL, Henderson Jr. CR, Cory-Slechta DA, et al. Intellectual impairment in children with blood Lead concentrations below 10 micrograms per deciliter. New Engl J Med 2003;348:1517-26.

Canfield RL, Gendle MH, Cory-Slechta DA. 2004. Impaired neurophysiological functioning in lead-exposed children. Developmental Neuropsychology 2004;26:513-40.

Chan TJH, Gutierrez C, Ogunseitan OA. Metallic burden of deciduous teeth and childhood behavioral deficits. Int J Environ Res Public Health 2015;12:6771-87.

Chandramouli K, Steer CD, Ellis M, et al. Effects of early childhood lead exposure on academic performance and behaviour of school age children. Archives of Disease in Childhood, 2009;94:844-8.

Choi WJ, Kwon HJ, Lim MH, et al. Blood lead, parental marital status and the risk of attention-deficit/hyperactivity disorder in elementary school children: a longitudinal study. Psychiatry Research 2016;236:42-6.

Claus Henn B, Schnaas L, Ettinger A, et al. 2012. Associations of early childhood manganese and lead co-exposure with neurodevelopment. Environ Health Perspect 2012;120:126-31.

Cooney GH, Bell A, McBride W et al. Low-level exposures to lead; the Sydney lead study. Developmental Medicine Child Neurology 1989;31:640-9.

Desrochers-Couture M, Oulhote Y, Arbuckle TE, et al. Prenatal, concurrent, and sex-specific associations between blood lead concentrations and IQ in preschool Canadian children. Environ Int 2018;121:1235-42.

Desrochers-Couture M, Courtemanche Y, Forget-Dubois N, et al. Association between early lead exposure and externalizing behaviors in adolescence: A developmental cascade. Environ Research 2019;178:108679.

Dietrich KN, Berger OG, Succop PA, et al. The developmental consequences of low to moderate prenatal and postnatal lead exposure: Intellectual attainment in the Cincinnati Lead study cohort following school entry. Neurotoxicol Teratol 1993;15:37-44.

Dikme G, Arvas A, Gur E. The relation between blood lead and mercury levels and chronic neurological diseases in children. Turk Arch Ped 2013;221-5.

Ernhart CB, Morrow-Tlucak M, Wolf AW, et al. Low level lead exposure in the prenatal and early preschool periods: Intelligence prior to school entry. Neurotox Teratol 1989;11:161-70.

Ethier AA, Muckle G, Jacobson SW, et al. Assessing new dimensions of attentional functions in children prenatally exposed to environmental contaminants using an adapted Posner paradigm. Neurotoxicol Teratol 2015;51:27-34.

Evens A, Hryhorczuk D, Lanphear BP, et al. The impact of low-level lead toxicity on school performance among children in the Chicago public schools: a population-based retrospective cohort study. Environ Health 2015;14:1-9.

Fergusson DM, Fergusson JE, Horwood LJ, et al. A longitudinal study of dentine lead levels, intelligence, school performance and behavior: part II: dentine lead and cognitive ability. J Child Psychology Psychiatry 1988;29:793-809.

Fergusson DM, Horwood LJ, Lynskey MT. Early dentine lead levels and subsequent cognitive and behavioural development. Journal of Child Psychology and Psychiatry, 1993;34:215-27.

Fergusson DM, Horwood LJ, Lynskey MT. Early dentine lead levels and educational outcomes at 18 years. Journal of Child Psychology and Psychiatry, 1997;38:471-8.

Forns J, Fort M, Casas M, et al. Exposure to metals during pregnancy and neuropsychological development at the age of 4 years. Neurotoxicol 2014;40:16-22.

Fruh V, Rifas-Shiman SL, Amarasiriwardena C, et al. Prenatal lead exposure and childhood executive function and behavioral difficulties in project viva. Neurotoxicol 2019;75:105-15.

Ha M, Kwon JH, Lim MH, et al. Low blood levels of lead and mercury and symptoms of attention deficit hyperactivity in children: A report of the children's health and environment research (CHEER). Neurotoxicol 2009;30:31-6.

Hong SB, Im MH, Kim JW, et al. Environmental lead exposure and attention deficit/hyperactivity disorder symptom domains in a community sample of South Korean school-age children. Environ Health Perspect 2015;123:271-6.

Huang PC, Su PH, Chen HY, et al. Childhood blood lead levels and intellectual development after ban of leaded gasoline in Taiwan: A 9-year prospective study. Environment International 2012;40:88-96.

Huang S, Hu H, Sanchez BN, et al. Childhood blood lead levels and symptoms of attention deficit hyperactivity disorder (ADHD): A cross-sectional study of Mexican children. Environ Health Perspect 2016;124:868-74.

Jedrychowski W, Perera F, Jankowski J, et al. Prenatal low-level lead exposure and developmental delay of infants at age 6 months (Krakow inner city study). International J Hygiene Environmental Health 2008;211:345-51.

Jedrychowski W, Perera FP, Jankowski J, et al. Very low prenatal exposure to lead and mental development of children in infancy and early childhood. Neuroepidemiol 2009;32:270-8.

Jedrychowski W, Perera F, Jankowski J, et al. Gender specific differences in neurodevelopmental effects of prenatal exposure to very low-lead levels: the prospective cohort study in three-year olds. Early Human Development, 2009;85:503-10.

Ji Y, Hong X, Wang G, et al. A prospective birth cohort study on early childhood lead levels and attention deficit hyperactivity disorder: new insight on sex differences. J Pediatrics 2018;199:124-31.

Joo H, Lim MH, Ha M, et al. Secondhand smoke exposure and low blood lead levels in association with attention-deficit hyperactivity disorder and its symptom domain in children: A community-based case-control study. Nicotine Tobacco Research 2017;94-101.

Joo H, Choi JH, Burm E, et al. Gender difference in the effects of lead exposure at different time windows on neurobehavioral development in 5-year-old children. Sci Tot Environ 2018;615:1086-92.

Jusko TA, Henderson Jr. C R, Lanphear BP, et al. Blood lead concentrations < 10 micrograms per deciliter and child intelligence at 6 years of age. Environ Health Perspect 2008;116:243-8.

Kim S, Arora M, Fernandez C, et al. Lead, mercury and cadmium exposure and attention deficit hyperactivity disorder in children. Environ Res 2013;126:105-10.

Kim KN, Kown HJ, and Hong YC. Low-level lead exposure and autistic behaviors in school-age children. NeuroToxicol 2016;53:193-200.

Lanphear BP, Hornung R, Khoury J, et al. Low-level environmental lead exposure and children's intellectual function: An international pooled analysis. Environ Health Perspect 2005;113:894-9.

Lee MJ, Chou MC, Chou WJ, et al. Heavy metals' effect on susceptibility to attention-deficit/hyperactivity disorder: Implication of lead, cadmium and antimony. Int J Environ Research Public Health 2018;15:1221.

Leviton A, Bellinger D, Allred EN, et al. Pre- and postnatal low-level lead exposure and children's dysfunction in school. Environmental Research 1993;60:30-43.

Lin Y, Huang L, Xu J, et al. Blood lead, bone lead and child attention-deficit hyperactivity disorder-like behavior. Sci Total Environ 2019;659:161-7.

Liu J, Li L, Wang Y, et al. 2013. Impact of low blood lead concentrations on IQ and school performance in Chinese children. PLoS One 2013;8(5)p.e65230.

Liu J, Liu X, Wang W, et al. Blood lead levels and children's behavioral and emotional problems: a cohort study. JAMA Pediatrics 2014;168:737-45.

Liu JA, Chen Y, Gao D, et al. Prenatal and postnatal lead exposure and cognitive development of infants followed over the first three years of life: a prospective birth study in the Pearl River Delta region, China. Neurotoxicology 2014;44:326-34.

Mazumdar M, Bellinger DC, Gregas M, et al. Low-level environmental lead exposure in childhood and adult intellectual function: a follow-up study. Environmental Health, 2011;10(1):24.

McDermott S, Wu J, Cai B, et al. Probability of intellectual disability is associated with soil concentrations of arsenic and lead. Chemosphere, 2011;84:31-8.

Min MO, Singer LT, Kirchner HL, et al. Cognitive development and low-level lead exposure in poly-drug exposed children. Neurotoxicology and Teratology 2009;31:225-31.

Miranda ML, Kim D, Galeano MAO, et al. The relationship between early childhood blood lead levels and performance on end-of-grade tests. Environ Health Perspect 2007;115:1242-7.

Miranda ML, Kim D, Reiter J, et al. Environmental contributors to the achievement gap. Neurotoxicology 2009;30:1019-24.

Moodie S, Ialongo N, Lopez P, et al. The conjoint influence of home enriched environment and lead exposure on children's cognition and behaviour in a Mexican lead smelter community. NeuroToxicol 2013;34:33-41.

Naicker N, Richter L, Mathee A, et al. Environmental lead exposure and socio-behavioural adjustment in the early teens: the birth to twenty cohort. Sci Total Environ 2012;414:120-5.

Neugebauer J, Wittsiepe J, Kasper-Sonnenberg M, et al. The influence of low level pre- and perinatal exposure to PCDD/Fs, PCBs, and lead on attention performance and attention-related behavior among

German school-aged children: Results from the Duisburg Birth Cohort Study. Int J Hygiene Environ Health 2015;218:153-62.

Nigg JT, Knottnerus GM, Martel MM, et al. Low blood lead levels associated with clinically diagnosed attention-deficit/hyperactivity disorder and mediated by weak cognitive control. Biol Psychiatry 2008;63:325-31.

Parajuli RP, Fujiwara T, Umezaki M, et al. 2013. Association of cord blood levels of lead, arsenic, and zinc with neurodevelopmental indicators in newborns: a birth cohort study in Chitwan Valley, Nepal. Environmental Research 2012;121:45-51.

Park JH, Seo JH, Hong YS, et al. Blood lead concentrations and attention deficit hyperactivity disorder in Korean children: a hospital-based case control study. BMC Pediatrics 2016;16:156.

Plusquellec P, Muckle G, Dewailly E, et al. The relation of low-level prenatal lead exposure to behavioral indicators of attention in Inuit infants in Arctic Quebec. Neurotoxicology Teratology 2007;29:527-37.

Plusquellec P, Muckle G, Dewailly E, et al. The relation of environmental contaminants exposure to behavioral indicators in Inuit preschoolers in Arctic Quebec. Neurotoxicology 2010;31:17-25.

Ris MD, Dietrich KN, Succop PA, et al. Early exposure to lead and neuropsychological outcome in adolescence. J International Neuropsychological Society 2004;10:261-70.

Rodrigues E, Bellinger D, Valeri L, et al. 2016. Neurodevelopmental outcomes among 2- to 3-year-old Children in Bangladesh with elevated blood lead and exposure to arsenic and manganese in drinking water. Environmental Health 2016;15:44.

Ruiz-Castell M, Paco P, Barbieri FL, et al. Child neurodevelopment in a Bolivian mining city. Environmental Research 2012;112:147-54.

Schnaas L, Rothenberg SJ, Perroni E, et al. 2000. Temporal pattern in the effect of postnatal blood lead level on intellectual development of young children. Neurotoxicology Teratology 2000;22:805-10.

Schnaas L, Rothenberg SJ, Flores MF, et al. Reduced intellectual development in children with prenatal lead exposure. Environ Health Perspect 2006;114:791-7.

Sioen I, Hond ED, Nelen V, et al. Prenatal exposure to environmental contaminants and behavioral problems at age 7-8 years. Environ Int 2013;59:225-31.

Taylor CM, Humphriss R, Hall A, et al. Balance ability in 7- and 10- year-old children: associations with prenatal lead and cadmium exposure and with blood lead levels in childhood in a prospective birth cohort study. BMJ 2015;5:1-8.

Taylor CM, Kordas K, Golding J, et al. Effects of low-level prenatal lead exposure on child IQ at 4 and 8 years in a UK birth cohort study. Neurotoxicology 2017;62:162-9.

Téllez-Rojo MM, Bellinger DC, Arroyo-Quiroz C, et al. Longitudinal associations between blood lead concentrations lower than 10 µg/dL and neurobehavioral development in environmentally exposed children in Mexico City. Pediatrics 2006;118:e323-e330.

Vigeh M, Yokoyama K, Matsukawa T, et al. Low level prenatal blood lead adversely affects early childhood mental development. Journal of child neurology, 2014;29:1305-11.

Wasserman GA, Liu X, Lolacono NJ, et al. Lead exposure and intelligence in 7-year-old children: the Yugoslavia prospective cohort study. Environ Health Perspect 1997;105:956-62.

Wasserman GA, Liu X, Popovac D, et al. The Yugoslavia Prospective Lead Study: contributions of prenatal and postnatal lead exposure to early intelligence. Neurotoxicology Teratology, 2000;22:811-8.

Winter AS and Sampson RJ. From lead exposure in early childhood to adolescent health: A Chicago birth cohort. Am J Public Health 2017;107:1496-501.

Wright JP, Dietrich KN, Ris MD, et al. Association of prenatal and childhood blood lead concentrations with criminal arrests in early adulthood. PLoS Medicine 2008;5(5):e101.

Yu CH, Du JC, Chiou HC, et al. Attention deficit/hyperactivity disorder and urinary nonylphenol levels: A case control study in Taiwanese children. PLoS One 2016;doi:10.1371.

Zhang R, Huo X, Ho G, et al. Attention-deficit/hyperactivity symptoms in preschool children from an E-waste recycling town: assessment by the parent report derived from DSM-IV. BMC Peds 2015;15:51.

Section 8.3 Other Documents

- Class Plaintiffs' Motion for Class Certification. (June 30, 2020). In re Flint Water Cases, Civil Action No. 5:16-cv-10444-JEL-MKM (consolidated)
- Declaration of Alan Ducatman, M.D., M.S. in support of plaintiffs' motion for class certification. In re Flint Water Cases, Civil Action No. 5:16-cv-10444-JEL-MKM (consolidated). (June 28, 2020).
- Deposition of Alan Ducatman, M.D. in support of plaintiffs' motion for class certification. In re Flint Water Cases, Civil Action No. 5:16-cv-10444-JEL-MKM (consolidated). (November 12, 2020 and December 7, 2020).
- Declaration of Alan Ducatman, M.D. In re: Michele Baker et al., Plaintiffs, v. Saint-Gobain Performance Plastics Corp., et al. Defendants. CIV. No. 1:16-cv-917 (LEK/DJS). In the United States District Court, Northern District of New York. (Filed April 6, 2020)
- Declaration of Howard Hu, M.D., M.P.H., Sc.D., in support of plaintiffs' motion for class certification. In re Flint Water Cases, Civil Action No. 5:16-cv-10444-JEL-MKM (consolidated). (June 29, 2020).
- Declaration of Daniel P. Keating, Ph.D., in support of plaintiffs' motion for class certification. In re Flint Water Cases, Civil Action No. 5:16-cv-10444-JEL-MKM (consolidated). (June 29 2020).
- Deposition of Daniel P. Keating, Ph.D., in support of plaintiffs' motion for class certification. In re Flint Water Cases, Civil Action No. 5:16-cv-10444-JEL-MKM (consolidated). (October 29, 2020 and December 3, 2020).

- Declaration of Bruce P. Lanphear, M.D., M.P.H., in support of plaintiffs' motion for class certification. In re Flint Water Cases, Civil Action No. 5:16-cv-10444-JEL-MKM (consolidated). (June 19, 2020).
- Declaration of Bruce P. Lanphear, M.D., M.P.H. In: Concerned Pastors for Social Action, et al., Plaintiffs, v. Nick A. Khouri, et al., Defendants. Case No. 16-10277. In the United States District Court for the Eastern District of Michigan, Southern Division. (March 23, 2016).
- Deposition of Bruce P. Lanphear, M.D., M.P.H. In: Concerned Pastors for Social Action, et al., Plaintiffs, v. Nick A. Khouri, et al., Defendants. Case No. 16-10277. In the United States District Court for the Eastern District of Michigan, Southern Division. (July 14, 2016).
- Videotaped Deposition of Bruce P. Lanphear, M.D. September 17 and 18, 2020.
- (Rough) Deposition of Howard Hu, M.D. October 12, 2020 and Final Corrected Deposition, November 5, 2020.
- Confidentiality Addendum to Plaintiff Darrell Davis's Response to Defendants' First Set of "Uniform" Interrogatories.
- Confidentiality Addendum to Plaintiff Darnella Gaines, as the Next Friend of K.C., a Minor, Response to VNA Defendants' First Set of Interrogatories to K.C., a Minor. September 9, 2020.
- Restricted Distribution—Confidential—KKelso—MDHHS—540006-000036-000250.
- Restricted Distribution—Confidential—RKelso—AGDHC—540007-000074.
- Restricted Distribution—Confidential—DDavis-QD-MD-540002-000009.
- Restricted Distribution—Confidential—DMunoz—GBCMC—540003-000002-000018.
- Restricted Distribuiton—Confidential—TWilliams—Henley MC—MD—540008-003128.
- ELNORA-CARTHAN_0000001.
- Tiantha-Williams 0000001.

Section 8.4 Websites

www.census.gov/quickfacts/fact/table/Geneseecounty
www.census.gov/quickfacts/fact/table/flintcitymichigan,MI/IPE120218.
www.mdch.state.mi.us) (accessed April, 2020)
www.michigan.gov/eatsafefish (accessed 3.25.2020)
www.cityofflint.com/lbphc (accessed 3.30.2020)
www.epa.gov/lead/learn-about-lead (accessed 3.24.2020)
(www.epa.gov/mercury (accessed 4.24.2020)
www.cdc.gov/nchs/nhanes (accessed 3.28.2020)
www.cdc.gov/nceh/lead/prevention/blood-lead-levels.htm (accessed 3.28.2020)
https://www.epa.gov/pcbs/learn-about-polychlorinated-biphenyls-pcbs (accessed 4.20.2020)
www.pehsu.net/_Library/facts/medical-mgmnt-childhood-lead-exposure-June-2013.pdf (accessed 12.19.2020)
www.nimh.nih.gov

PART NINE: APPENDICES

Appendix A Systematic Assessments of Non-Lead Causes of Neurodevelopmental Outcomes in Children (i.e. the "Other Contaminants")

Literature Search and Included/Excluded Published Papers

Systematic searches of the biomedical literature were performed for English-language reviews, systematic reviews, and meta-analyses on the topics mentioned above.

Medline (PubMed) was searched for English-language publications as follows:

PubMed (November 18, 2019): search terms "smoking" and "neurodevelopment" + "children" (n=19).
PubMed (November 18, 2019): search terms "obesity" and "neurodevelopment" + "children" (n=36).
PubMed (March 2, 2020): search terms "phthalates" and "neurodevelopment" (n=51).
PubMed (November 19, 2019): search terms "air pollution" and "neurodevelopment" + "children" (n=16).
PubMed (January 22, 2020): search terms "pesticides" and "neurodevelopment" + "children" (n=63).
PubMed (March 16, 2020): search terms "polybrominated diphenyl ethers" and "neurodevelopment" and "children" (n=24).
PubMed (April 25, 2020): search terms "mercury" and "neurodevelopment" and "children" (n=128).
PubMed (April 22, 2020): search terms "preterm" and "neurodevelopment" and "children" (n=
PubMed (May 6, 2020): search terms "polychlorinated biphenyls" and "children" and "neurodevelopment" (n=68).
PubMed (October 14, 2020): search terms "fluoride" and "children" and "neurodevelopment" (n=88).

Because these searches all involve the terms "neurodevelopment" and "children," there are likely duplicates. After removal of duplicates, reviews, systematic reviews, and meta-analyses were identified by topic as described below.

Recent Reviews and More Recent Studies of Smoking and Neurodevelopment in Children

Reviews of Smoking and Neurodevelopment in Children: Clifford et al. (2012), Chen et al. (2013), and Polanska et al. (2015).

Clifford et al. (2012) is a narrative review using a structured systematic approach. The purpose of the review was to identify the relevant literature and "examine whether recently published studies have demonstrated a relationship between exposure to tobacco smoke in utero and cognitive functioning…" (Clifford et al. 2012, p. 561). The authors identified "observational studies published between 2000 and 2011 that examined associations between tobacco smoke exposure in utero due to maternal smoking

and performance on cognitive, intelligence, neurodevelopmental and academic tests" (Clifford et al. 2012, p. 560). Twenty publications were selected for analysis, representing 18 studies. Smoking by the mothers was assessed primarily by self-report, retrospectively in 4 studies and prospectively in the remaining studies. The review examined the following neurodevelopmental outcomes: intelligence (6 prospective and 2 retrospective studies), intellectual impairment—i.e. IQ score below 70 points—(2 prospective studies), memory (5 studies), attention and executive function (7 studies), global neurodevelopment (1 study), and academic achievement (3 studies).

The authors concluded the following (Clifford et al. 2012, p. 560):

> "This review found evidence of a relationship between tobacco smoke exposure in utero and reduced academic achievement and cognitive abilities independent of other variables."

Chen et al. (2013) is a systematic narrative review of publications (1989-2012) investigating the association between secondhand smoke (SHS) exposure and performance on neurocognitive and academic tests. Studies examining in utero SHS exposure by pregnant women and studies examining postnatal exposure to SHS were included. The authors noted that they focused on SHS not due to maternal smoking during pregnancy given the **Clifford et al. (2012)** review described immediately above. Studies included were those with a clear measure of SHS and at least one objective measure of cognitive performance; in addition, studies were excluded if they did not take account of the effects of other toxins such as urban pollutants or illicit drug exposure. Risk of bias was assessed. Fifteen (15) studies were included in the analyses, representing four (4) cross-sectional studies and eleven (11) prospective studies. Five studies used cotinine level as an objective measure of SHS exposure whereas the remainder relied upon self-report. The review examined the following neurodevelopmental outcomes: cognitive outcomes (7 prospective studies), SHS exposure and cognition in preschool children (4 prospective studies), and SHS exposure and cognition in older children (4 cross-sectional studies) and (2 prospective studies).

The authors concluded the following (Chen et al. 2013, p. 2013):

> "Prenatal SHS exposure was inversely associated with neurodevelopmental outcomes in young children, whereas postnatal SHS exposure was associated with poor academic achievement and neurocognitive performance in older children and adolescents."

In addition, these authors conclude (Chen et al. 2013, p. 660):

> "Overall, SHS exposure in utero appears important to global cognitive functioning and development over the first few years of life, whereas postnatal SHS exposure seems to become important later in childhood."
>
> "Based on the findings of our systematic literature review, further campaigns aimed at discouraging cigarette smoking and avoiding SHS exposure could contribute to the prevention of cognitive impairment, slowing the trend of epidemic dementia worldwide."

Polanska et al. (2015) is a narrative review of English language epidemiological studies published between 2006 and 2013 that examined the relationships between

neurodevelopmental outcomes and prenatal exposure to tobacco smoke or low or moderate alcohol drinking. Both exposure to active and passive smoking studies were included. Excluded were studies examining fetal alcohol syndrome, binge and heavy drinking studies, studies of postnatal environmental tobacco smoke, combination of exposure studies (i.e. studies that combined different types of toxicants into a single variable) as well as review papers and papers discussing mechanism.

Studies of Smoking and Neurodevelopmental Outcomes

The authors write that they assume approximately 20-30% of women actively smoke during pregnancy and about half of non-smoking pregnant women are exposed to passive smoking, citing a report of the U.S. Surgeon General (2010). In addition, the authors assume that "almost half of the child population is involuntarily exposed to tobacco smoke at home, citing the American Cancer Society (2012). The authors note that more studies examined the impact of prenatal exposure to active maternal smoking on cognitive development, intelligence, and intellectual impairment, describing 7 studies. They write that "some of them indicated the association between such exposure and decreased IQ or cognitive delay, there is controversy whether possible cofounders have been adequately controlled for given the strong influences of parental IQ, socioeconomic status, home environment, and sibling relationships.

The authors conclude that the results of these studies "are not fully consistent" (Polanska et al. 2015, p. 429).

Studies of Low-to-Moderate Alcohol Consumption and Neurodevelopmental Outcomes

The authors examined studies in which low consumption of alcohol was defined as 1-2 drinks per week or less than 4 drinks per week, whereas moderate drinking was defined as somewhere between 3-6 or 5-8 drinks per week. The authors examined approximately 11 studies and concluded that the results were too inconsistent and thus only "suggest that consumption of alcohol during pregnancy may adversely affect children's intelligence quotient (IQ), mental health, memory and verbal or visual performance."

More Recent Studies of Smoking and/or Alcohol and Neurodevelopment in Children

Because the most recent published review on smoking and/or alcohol use pre- and postnatally examined studies through 2013, it is important to examine the epidemiological studies published between 2014 and the present (i.e. January, 2020).

A systematic search of PubMed was undertaken on January 20, 2020 using the search terms "smoking" and "neurodevelopment" and "children" to include English-language epidemiological studies. There were 62 publications identified, of which the following were considered relevant after full-text review: Moore et al. (2019), Roige-Castellvi et al. (2019), Furlong et al. (2018), Massey et al. (2018), Polanska et al. (2017), Wiebe et al. (2015), Evlampidou et al. (2015), Han et al. (2015), Polanska et al. (2015). Of these, Massey et al. (2018) was excluded because neurodevelopmental outcomes were not included.

Brief descriptions of the conclusions of these studies follow.

Evlampidou et al. (2015, p. 246) conclude:

> "Maternal exposure during pregnancy to second-hand smoke measured through urine cotinine was associated with a decrease in gross motor function among 18-month-old children, even at low levels of exposure."

Han et al. (2015, p. 164) conclude:

> "…this study is meaningful because it identified an increase in the prevalence of ADHD…"

Kesmodel et al. (2014, p. 1736) conclude:

> "In summary, we observed that the intake of ≥22 drinks/week on average before pregnancy was associated with lower mean full-scale IQ, overall attention and sustained attention but not with selective attention, executive function or motor function."

Moore et al. (2019, p. 4) conclude:

> "Our results confirm earlier findings of less optimal neurocognitive development among children exposed to tobacco in utero. In addition, we provide evidence that early life exposure to tobacco smoke is associated with less optimal fine motor development and reduced inhibitory control in children born at ≥37 weeks of gestation and birth weights of ≥2500 g."

Polanska et al. (2015, p. 317) conclude:

> "Children prenatally exposed to tobacco compounds and those of underweight mothers had a decreased psychomotor development."

Polanska et al. (2017, p. 8 of 12) conclude:

> "The present study showed that 30% of the non-smoking women were exposed to ETS during the pregnancy period. Children of the mothers who were passive smokers while pregnant had decreased neurodevelopment abilities compared with children of the non-exposed mothers. This association was observed for the whole spectrum of child neurodevelopment, including cognitive, language, and motor abilities."

Wiebe et al. (2015, p. 404-5) conclude:

> "In this study, we examined whether smoking during pregnancy was related to self-regulation in early childhood. We predicted that children with PTE (i.e. prenatal tobacco exposure) would show lower levels of self-regulation in both cognitive and motivational domains; this prediction was only partially supported, because PTE-related differences were present only in the motivational domain. We also predicted that the relation between PTE and self-regulation would be stronger in boys than in girls. This hypothesis was supported, because PTE's impact on motivational self-regulation was significantly greater for boys than for girls."
>
> Note that the effect in boys remained after adjusting for covariates.

Summary of Studies Published After Recent Reviews: Smoking and Neurodevelopment in Children

These studies add to the consistency of studies examining the relationship between exposure to cigarette smoke during pregnancy and a number of neurodevelopmental outcomes, including gross motor function, ADHD, fine motor development, inhibitory control, psychomotor development, cognitive, language and motor abilities, and self-regulation.

Recent Reviews and More Recent Studies of Maternal Weight (i.e.Overweight/Obesity) and Neurodevelopment in Children

Reviews of Overweight/Obesity and Neurodevelopment in Children: Van der Burg et al. (2016), Alvarez-Bueno et al. (2017), and Sanchez et al. (2018).

Van der Burg et al. (2016) is a narrative review of the mechanistic hypothesis that maternal obesity causes maternal and fetal inflammation and that this inflammation adversely affects the neurodevelopment of children. The authors included a table of epidemiological studies dated between 2004 and 2013 from which they conclude the following (Van der Burg et al., 2016, p. 2):

> "Children of mothers who were overweight or obese during pregnancy were at elevated risk of four major categories of neurodevelopmental deficits, including cognitive and intelligence deficits, attention deficit hyperactivity disorder (ADHD), autism, and psychoses."

Alvarez-Bueno et al. (2017) is a systematic review and meta-analysis of epidemiological studies published through February, 2017 examining the relationship between pre-pregnancy weight status and children's cognition. Inclusion criteria were clearly stated. Selected were follow-up studies involving pre-pregnancy weight status and children's cognition, with calculated pre-pregnancy weight status, and children's cognition assessed by standardized test scores or curricular-based grades related to specific subject areas. Exclusions: studies not published in English or Spanish, studies of women with intellectual disabilities, children not born at full term, and children with mental disorders that could limit generalizability [ADHD, conduct or neuropsychiatric disorders, or detected delay in communication, adaptive, cognition, or socio-emotional domains]. In other words, the authors examined studies of "healthy" children not known to have neurocognitive deficits at birth. Study quality was assessed using a published checklist, consisting of 33 items. The authors calculated the effect size (ES), i.e. a measure of change in the outcome variable, using the Mantel-Haenszel fixed effects model. 95% confidence intervals were also calculated as were heterogeneity statistics using I^2. The authors identified 15 studies that met their inclusion criteria; 9 of those studies were included in the meta-analyses. Results of the meta-analysis examining neurocognitive outcomes were as follows:

Neurocognitive Outcomes and Pre-Pregnancy Weight Status

Pre-Pregnancy Status	ES (95% CI)	I^2
Overweight	-0.02 (-.0.05 to 0.02)	0%
Obese	-0.06 (-0.09 to -0.03)	0%
Excess Weight	-0.04 (-0.06 to -0.02)	0%

The authors conclude that "pre-pregnancy obesity might have negative consequences on the neurocognitive development of offspring" (Alvarez-Bueno et al. 2017, p. 1653).

Results of the meta-analysis examining children's intelligence were as follows:

Intelligence Outcomes and Pre-Pregnancy Weight Status

Pre-Pregnancy Status	ES (95% CI)	I^2
Overweight	-0.02 (-.0.06 to 0.02)	0%
Obese	-0.05 (-0.10 to 0.00)	0%
Excess Weight[1]	-0.03 (-0.06 to 0.00)	0%

[1]Excess Weight = Combines categories of overweight and obesity

The authors conclude that "pre-pregnancy obesity could have a small negative, but not statistically significant, effect on children's general intelligence" (Alvarez-Bueno et al. 2017, p. 1660).

Sanchez et al. (2018) is the report of a systematic review and meta-analysis of epidemiological studies of pre-pregnancy overweight/obesity status and child neurodevelopmental outcomes. The authors identified studies published through April, 2017. Of the 41 studies identified using systematic search techniques, 32 publications representing 36 cohorts were included in the meta-analyses. Study quality was assessed using published scales; quality indicators included type of study, loss to follow-up, sample size, participant selection, comparability of groups, statistical methods, and criteria for determining and categorizing pre-pregnancy weight (measured as BMI) and neurodevelopmental outcomes. Studies were considered of high, medium, or low quality. The meta-analysis examined odds ratios (ORs) and 95% confidence intervals (95% CI) from each study, with an emphasis on adjusted vs. unadjusted measures. A random effects model was used to account for possible heterogeneity—also evaluated using the statistic, I^2—between studies. Publication bias and sensitivity analyses were performed.

Results are shown in the tables below, adapted from Sanchez et al. (2018, Table 2, p. 476):

Risk of Any Adverse Neurodevelopmental Outcome by Weight Group

BMI Weight Group	# Cohorts	OR (95% CI)	p-value	I^2
Overweight	22	1.17 (1.11-1.24)	<0.001	66.51
Obese	25	1.51 (1.35-1.69)	<0.001	79.63
Excess Weight[1]	29	1.33 (1.22-1.45)	<0.001	93.07

[1]Excess Weight = Combines categories of overweight and obesity

Risk of ADHD by Weight Group

BMI Weight Group	# Cohorts	OR (95% CI)	p-value	I^2
Overweight	6	1.30 (1.10-1.54)	0.01	52.97
Obese	7	1.62 (1.23-2.14)	0.015	60.52
Excess Weight[1]	8	1.51 (1.28-1.77)	<0.001	82.11

[1]Excess Weight = Combines categories of overweight and obesity

Risk of ASD (Autism Spectrum Disorder) by Weight Group

BMI Weight Group	# Cohorts	OR (95% CI)	p-value	I^2
Overweight	10	1.10 (1.01-1.21)	0.026	0.00

Obese	11	1.36 (1.08-1.70)	0.015	60.52
Excess Weight[1]	11	1.23 (1.06-1.42)	0.013	77.36

[1]Excess Weight = Combines categories of overweight and obesity

Risk of Cognitive/Intellectual Delay by Weight Group

BMI Weight Group	# Cohorts	OR (95% CI)	p-value	I^2
Overweight	11	1.19 (1.09-1.29)	0.003	40.05
Obese	14	1.58 (1.39-1.79)	<0.001	75.77
Excess Weight[1]	17	1. 35 (1.15-1.57)	<0.001	94.03

[1]Excess Weight = Combines categories of overweight and obesity

Risk of Emotional/Behavioral Problems by Weight Group

BMI Weight Group	# Cohorts	OR (95% CI)	p-value	I^2
Overweight	7	1.14 (0.93-1.39)	0.16	70.78
Obese	6	1.42 (1.26-1.59)	<0.001	87.74
Excess Weight[1]	10	1.51 (1.28-1.77)	<0.001	90.37

[1]Excess Weight = Combines categories of overweight and obesity

The authors conclude from these data "show that children born to mothers who are overweight or obese are at higher risk of neurodevelopmental problems, including ADHD, ASD, greater emotional/behavioral problems and cognitive delay" (Sanchez et al. 2018, p. 476).

<u>More Recent Studies of Maternal Weight (Overweight/Obesity) and Neurodevelopment in Children</u>

Because the most recent published review on weight (obesity/overweight) and neurodevelopmental outcomes evaluated studies published through mid-2017, it is important to examine the epidemiological studies published between mid-2017 and the present (i.e. January, 2020).

A systematic search of PubMed was undertaken on January 21, 2020 using the search terms "obesity" and "weight" and "neurodevelopment" and "children" to include English-language epidemiological studies. There were 20 publications identified, of which the following were considered relevant after full-text review: Krzeczkowski et al. (2019), Nichols et al. (2019), Widen et al. (2019), Widen et al. (2018), and Yeung et al. (2017).

Brief descriptions of the conclusions of these studies follow.

Krzeczkowski et al. (2019 p. 8) conclude:

> "In the adjusted models, maternal metabolic complications (including obesity) were not associated with offspring neurodevelopment."

Nichols et al. (2019, p. 55) conclude:

> "In our low-income, minority cohort, pre-pregnancy obesity was associated with lower PDI (psychomotor development) scores at 3 years only in boys."

Widen et al. (2019, p. 1) conclude:

> "Prepregnancy overweight and obesity were associated with lower IQ among boys, but not girls, at 7 years."

Widen et al. (2017, p. 6) conclude:

> "In summary, we found that maternal pre-pregnancy overweight and obesity were associated with a reduction in verbal recognition cognitive test scores in mid-childhood."

Yeung et al. (2017, p. 8) conclude:

> "In this first examination of maternal and paternal obesity in the United States on early childhood development, maternal obesity was associated with delays in fine motor development and paternal obesity marginally associated with delays in personal-social functioning. The impact of higher levels of parental obesity (i.e. having both parents with BMI ≥35, which constituted 3% of our cohort was most striking for multiple domains."

Summary of Studies Published After Recent Reviews: Overweight/Obesity and Neurodevelopment in Children

These studies add to the consistency of studies examining the relationship between exposure to maternal (and paternal) obesity during pregnancy and several neurodevelopmental outcomes.

Recent Reviews and More Recent Studies of Pesticide Exposure and Neurodevelopment in Children

Reviews of Pesticide Exposure and Neurodevelopment in Children: Koureas et al. (2012), Burns et al. (2013), Munoz-Quezada et al. (2013), Gonzalez-Alzaga et al. (2014), Hertz-Picciotto et al. (2018), and Tsai et al. (2019).

Koureas et al. (2012) is a systematic "mini-review" of epidemiological studies that used biomonitoring techniques to define exposure. As the authors note (Koureas et al. 2012, p. 156), "biomonitoring involves measurement of the parent compound or its metabolites in human biological samples in order to identify and quantify the internal exposure to specific chemicals." Exposures of interest in this review were the organophosphorous (OP) and pyrethoid (PYR) insecticides.

The authors identified studies published through 2011. Only one study had been published on PYR insecticides and neurodevelopment in children and it did not show an association between prenatal exposure and neurodevelopment. Studies of OP insecticides included prospective designs: the CHAMACOS study, the CCEH study, and a study from Mt. Sinai, as well as cross-sectional studies typically using NHANES data. Significant findings were reported between exposure and reflexes, cognitive abilities, attention problems, intellectual development, mental development, mental and motor delays, attention problems, attention/hyperactivity disorder, memory, and pervasive developmental disorders.

The authors conclude that "these studies produced significant evidence that women exposed to OP during pregnancy is correlated with neurodevelopmental effects in children" (Koureas et al. 2012, p. 163). The authors caution that some concerns have been raised regarding these findings and that "safe conclusions cannot be made."

Burns et al. (2013) is a systematic review of epidemiological and animal studies of pesticide exposure and neurodevelopmental outcomes. The authors identified 46 publications from 16 epidemiological studies that examined exposure to pesticides and head circumference and/or neurobehavioral endpoints. Two major categories of pesticides were the focus of these studies: organophosphate (OP) and organochlorine (OC). It should be noted that this is a complex and long—150 page—published review that concludes the following: "as a whole, the epidemiologic studies did not strongly implicate any particular pesticide as being causally related to adverse neurodevelopmental outcomes in infants and children. A few associations were unique for a health outcome and specific pesticide, and alternative hypotheses could not be ruled out" (Burns et al., 2013, p. 127). However, the authors do not clearly state which pesticides and which outcomes represent associations.

Munoz-Quezada et al. (2013) is a systematic review of epidemiological studies published between 2002 and 2012 (English or Spanish) of prenatal and early childhood exposure to organophosphate (OP) pesticides and neurodevelopmental effects in children. Twenty-seven (27) publications met the inclusion criteria, primarily undertaken in the U.S. (n = 16). Study quality was assessed based on exposure assessment, neurodevelopmental assessment, study design, sample size, and control of confounding. A score of 8-10 means "high quality;" a score of 3-7 means "intermediate rating;" and a score of 0-2 means "low rating." Prospective cohort studies were given higher scores than case-control studies. No score (0) was given to cross-sectional studies for the study design consideration. The authors did not perform a meta-analysis due to the many different outcomes and different analytical techniques. Nevertheless, the authors conclude that "all studies except one cross-sectional study in China provided evidence that exposure to OP pesticides was a risk factor for poor neurodevelopment, with the strongest linkages resulting from prenatal exposures" (Munoz-Quezada et al. 2013, p. 164). Furthermore, they write that the studies with the highest quality rankings—10 longitudinal studies—found positive dose-response relationships between OP exposure and neurodevelopmental outcomes, which "strengthens a presumption of causality" (Munoz-Quezada et al. 2013, p. 164). The outcomes of concern included cognitive, behavioral (primarily related to attention problems) and motor.

Gonzalez-Alzaga et al. (2014) is a systematic review of neurodevelopmental effects of prenatal and postnatal exposure to organophosphate pesticides. The authors identified studies published before or during December, 2012, written in English, Spanish, Portuguese, or French, evaluated prenatal or postnatal exposure to OP pesticides, and used general intelligence tests or other specific tests to assess changes in mental or motor development or behavior in children. Methodological quality of the studies was assessed using the STROBE checklist; studies were categorized as low, moderate, and high quality. Twenty studies were identified, 10 cohort studies, 9 cross-sectional studies, and 1 case-control study. The authors conclude that the studies identified and evaluated "suggest that exposure during pregnancy may have a negative effect on the child's mental and motor development and behavior during the first stages of childhood" (Gonzalez-Alzaga et al. 2014, p. 119). The outcomes of primary concern were mental development and attention problems in preschool and school children.

Hertz-Picciotto et al. (2018) is a policy-oriented commentary that discusses the overall evidence on exposure to organophosphate pesticides and neurodevelopmental outcomes in children, citing several recent reviews. The authors write (Hertz-Picciotto et al. 2018, p. 2):

> "Systematic reviews and multiple epidemiological studies in the US and other countries, spanning diverse populations in both urban and agricultural settings, have linked OP exposures during fetal development with poorer cognitive, behavioral, and social development in children."

Tsai et al. (2019) is a review of studies undertaken in Asia on children's health given exposure to air pollution, pesticides, and heavy metals. Only the authors' review of pesticide-related studies published through January 31, 2019 will be examined here. For neurodevelopmental and behavioral problems, the authors identified 5 studies each of which reported decreases in behavioral or cognitive function.

More Recent Studies of Pesticide Exposure and Neurodevelopment in Children

Because the most recent published review on pesticide exposure and neurodevelopmental outcomes evaluated studies published through late 2012, it is important to examine the epidemiological studies published between late 2012 and the present (i.e. January, 2020).

A systematic search of PubMed was undertaken on January 22, 2020 using the search terms "pesticide" and "neurodevelopment" and "children" to include English-language epidemiological studies. There were 63 publications identified, of which the following were considered relevant after full-text review: Cartier et al. (2016), Coker et al. (2017), Donauer et al. (2016), Engel et al. (2016), Fluegge et al. (2016), Furlong et al. (2017a), Furlong et al. (2017b), Gunier et al. (2017), Quiros-Alcala et al. (2014), Viel et al. (2015), Viel et al. (2017), and Zhang et al. (2014).

Brief descriptions of the conclusions of these studies follow.

Cartier et al. (2016, p. 679) conclude:

> "Using first-trimester urinary biomarkers of OP exposure and standardized cognitive tests to assess the potential developmental neurotoxicity of Ops in the European general population, we did not observe associations between biomarkers of prenatal OP exposures and children's cognitive scores at age 6 years, except that DE exposure appeared to be associated with better verbal comprehension scores."

Coker et al. (2017, p. 15) conclude:

> "We observed that pesticide-use profiles of neurotoxic pesticides used near the homes of pregnant women living in agricultural communities were associated with FSIQ deficits in their children."

Donauer et al. (2016, p. 414-5) conclude:

> "In this study, children of women exposed to organophosphate pesticides during pregnancy at levels that resemble average US levels as determined by the National Health and Nutrition Examination Survey (NHANES) were assessed for cognitive, motor, language, and intelligence outcomes from 1 to 5 years of age. In unadjusted analyses, several significant positive associations were found between urinary concentrations of

> metabolites of organophosphate pesticides, but these associations became insignificant in multivariable analyses that included adjustment for covariates of exposure and outcome. The positive associations found in the unadjusted analyses in this cohort are likely a result of confounding by socioeconomic status and dietary intake of fruits and vegetables, and not due to beneficial properties of pesticides."

Engel et al. (2016, p. 827) conclude:

> "In this pooled analysis of four prospectively enrolled NIEHS/EPA-funded Children's Environmental Health and Disease Prevention Research Centers, we estimated that each 10-fold increase in prenatal exposure to ΣDAPs is associated with an approximate 1-point decrease in the 24-month MDI (Mental Development Index), whether pooled across four cohorts without adjustment for race/ethnicity (β = -1.28; 95% CI:-2.58, 0.03) or pooled across race/ethnicities without adjustment for cohort (β = -1.48; 95% CI: -2.77, -0.19)."

Fluegge et al. (2016, p. 60) conclude:

> "These results were robust to second month infant urine measures of 3,5,6-trichloro-2-pyridinol (metabolite of OP chlorpyrifos), which independently had a significant and negative influence on mental functioning."

Furlong et al. (2017a, p. 2) introduce their study of prenatal exposure to organophosphate pesticides—including phthalates, phenols, and pyrethroids—and neurodevelopmental phenotypes by noting that prenatal exposure to organophosphate pesticides (OPs) has been associated with impaired neurodevelopment in both urban and agricultural populations. Specifically, prenatal OP exposure has been associated with measures of cognition, including lower IQ scores and lower scores on the Bayley Scales of Infant Development Mental Development Index as well as various measures of behavior, including impaired social responsiveness, indicators of Pervasive Developmental Disorder and inattention.

These authors (2017 p. 13) conclude:

> "Standardized dimethylphosphate metabolites (ΣDMPs) were negatively associated with Internalizing factor scores…"

> "Standardized diethylphosphate metabolites (ΣDEPs) were negatively associated with the Working Memory Index…."

Importantly, these authors also note that "other environmental exposures, such as flame retardants, organochlorine pesticides or PCBs, air pollution, lead, heavy metals, as well as others, may confound our associations. We did not measure flame retardants, air pollution, non-lead heavy metals in this population, and organochlorines, PCBs, and lead were only measured on a subset of the population. Including them would have eliminated approximately a third of our population and destabilized the analysis" (Furlong et al. 2017, p. 13).

Furlong et al. (2017b, p. 231) conclude:

"In longitudinal mixed models, detectable levels of 3-PBA (a metabolite of pyrethroid pesticides) during pregnancy were associated with worse internalizing, depression, somatization, behavioral regulation, emotional control, shifting, and monitoring."

"Detectable levels of cis-DCCA (a metabolite of pyrethroid pesticides) were associated with worse externalizing, conduct problems, behavioral regulation, and inhibitory control."

Gunier et al. (2017, p. 1) conclude:

"We observed a decrease of 2.2 points (95% confidence interval (CI): -3.9 to -0.5) in Full-Scale IQ and 2.9 points (95% CI: -4.4 to -1.3) in Verbal Comprehension for each standard deviation increase in toxicity-weighted use of organophosphate pesticides."

Quiros-Alcala et al. (2014, p. 1336) conclude:

"In conclusion, postnatal pyrethroid exposure was not significantly associated with parent-reported LD, ADHD, or both LD and ADHD in children 6-15 years of age participating in the NHANES during 1999-2002. However, the cross-sectional nature of this study does not allow us to establish causality not to examine exposure during the potentially more critical prenatal period."

Viel JF et al. (2015, p. 2) conclude:

"…childhood 3-PBA and cis-DBCA concentrations were both negatively associated with verbal comprehension scores (P-trend=0.04 and P-trend<0.01, respectively) and with working memory scores (P-trend=0.05 and P-trend<0.01, respectively)."

Viel JF et al. (2017, p. 2) conclude:

Increased prenatal cis-3-(2,2-dichlorovinyl)-2, 2-dimethylcyclopropane carboxylic acid (DCCA) concentrations were associated with internalizing difficulties (Cox p value = 0.05).For childhood 3-phenoxybenzoic acid (PBA) concentrations, a positive association was observed with externalizing difficulties (Cox pvalue=0.04) and high odds ratios (ORs) were found for abnormal or borderline social behavior (OR=2.93, 95% CI: 1.27, 6.78, an d OR=1.91, 95% CI: 0.804.57, for theintermediate and highest metabolite categories, res pectively). High childhood transDCCA concentrations were associated with reduced exte rnalizing disorders (Cox p-value=0.03). "

Zhang et al. (2014, p. 1) conclude:

"The high exposure of pregnant women to OPs in Shenyang, China was the predominant risk factor for neonatal neurobehavioral development."

Recent Reviews and More Recent Studies of Exposure to Phthalates and Neurodevelopment in Children

Reviews of Exposure to Phthalates and Neurodevelopment in Children: Jurewicz and Hanke (2011), Jurewicz et al. (2013), Ejaredar et al. (2015), Lee et al. (2018), Zhang et al. (2019), and Radke et al. (2020).

Jurewicz and Hanke (2011) is a narrative review of the epidemiological studies examining the relationship between exposure to phthalates and various outcomes in children including neurodevelopmental outcomes. The authors identified epidemiological studies published between 2000 and 2010. Four (4) studies of exposure to phthalates and neurodevelopmental outcomes were identified; see Jurewicz and Hanke (Table 4, p. 125) for a tabular description of those studies. The authors of the review conclude that "the findings of the reviewed studies indicate that children's exposure to phthalates (MBzP, MECPP, MEHHP, MEOHP, MEHP, MCPP, MEP, MBP, MiBP, ΣDEHP, ΣDBP) may bring about impairments in the neurodevelopmental processes" (Jurewicz and Hanke 2011, p. 127). Outcomes involved include poor scores on conduct problems, quality of alertness among girls, decreased composite scores (boys), and lower concentrations and vocabulary score among boys as well as attention deficit hyperactivity disorder (ADHD).

Jurewicz et al. (2013) is a narrative review of studies examining exposure to environmental toxicants and children's cognitive development and behavioral problems. Phthalates are one of those toxicants as well as bisphenol A, brominated flame retardants, polycyclic aromatic hydrocarbons (PAHs), as well as exposure to indoor gas cooking. Here, only the studies involving phthalates will be described and discussed. Studies published since 2000 through 2012 were included in the review; there were seven (7) studies of phthalates and neurodevelopmental outcomes included in the review. See Jurewicz et al. (Table 1, p. 188-9) for a description and discussion of these studies. The authors conclude that "the findings of the reviewed studies indicate that children's exposure to phthalates (MBzP, MECPP, MEHHP, MEOHP, MEHP, MCPP, MEP, MBP, MiBP, ΣDEHP, ΣDBP) may bring about impairments in the neurodevelopmental processes" (Jurewicz et al. 2013, p. 190).

Ejaredar et al. (2015) is a systematic review of studies examining relationships between exposure to phthalates and children's neurodevelopmental outcomes. The authors described that the review was undertaken using the MOOSE guidelines. Eleven (11) studies of exposure to phthalates and neurodevelopmental outcomes in children aged 0-12 years were identified using various databases; outcomes included measures of cognition, internalizing behaviors (e.g. anxiety, depression), and externalizing behaviors (e.g. aggression). In the studies included in the review, exposure to phthalates was defined as a quantitative measure of the concentration of phthalate metabolites in urine. Children's neurodevelopmental outcomes were measured using standardized cognitive and behavioral measures (e.g. Bayle Scales of Infant Development [BSID-11], and Behavioral Assessment System for Children-2 [BASC-2]. Results of these studies are summarized in Ejaredar et al. (Table 1, 2015); quality assessment of these studies can be found in Ejaredar et al. (Table 5, p. 56); the measures involved in the studies can be found in Ejaredar et al. (Table 6, p. 57); and the confounding factors adjusted for in the studies in Ejaredar et al. (Table 7, p. 58). The authors conclude (Ejaredar et al. 2015, p. 58):

> "This systematic review suggests that higher levels of prenatal exposure to phthalate metabolites measured as urinary concentrations are associated with poorer cognitive and behavioral outcomes in children 0-12 years of age."

Lee et al. (2018) is a systematic review and meta-analysis focused solely on a single phthalate—di-(2-ethylhexyl) phthalate (DEHP)—and neurodevelopmental outcomes. The authors searched for epidemiological studies published between January 1, 1980 and December 31, 2017 examining these relationships. Searching various databases, the authors identified English-language cohort, case-control, and cross-sectional studies; they excluded letters to the editor, commentaries, and reviews. Only the most recent publication of any study was included; earlier published analyses of the same study were excluded. The authors used the Newcastle-Ottawa scale (NOS) to evaluate study quality. Ten (10) studies were identified and included in the meta-analysis; outcome measures included the Wechsler Intelligence Scale (WISC) or the Bayley Scale of Infant Development (BSID). The authors concluded (Lee et al. 2018, p. 562):

> "We found a significant negative association between childhood exposure to environmental DEHP and neurodevelopmental outcome measured using the WISC. A 2-fold increase in DEHP metabolites of childhood is associated with a 0.8-point reduction in IQ. Furthermore, a 2-fold increase in DEHP metabolites of prenatal maternal urine is associated with a 0.6-point reduction in PDI of their childhood."

Zhang et al. (2019) is a review of the association between prenatal exposure to phthalates and cognitive and neurobehavioral outcomes. The focus of the review is a specific study design, namely, birth cohorts. Through searches of PubMed, EMBASE, and Web of Science, the authors identified twenty-six (26) birth cohort studies; nine (9) of the studies examined cognition, thirteen (13) examined neurobehavior, and four (4) examined both cognition and neurobehavior. The searches were undertaken in October, 2017 and May, 2018. Excluded were reviews, meta-analyses, meetings, letters, abstracts, and commentaries. The authors identified English-language studies of prospective birth cohorts in which phthalate metabolites were measured in urine, and the methods used were generally accepted. Each study was evaluated using the Cochrane Collaboration's "Risk of Bias" tool.

The authors (Zhang et al. 2019, p. 209 and p. 210) concluded the following:

> "In general, the results showed that prenatal exposure to DEHP, BBzP, DEP, and DBP influenced cognition and behavior development. Exposure to phthalates had adverse effects on cognitive development in boys and girls, with gender differences. In addition, phthalate exposure was more likely to cause behavioral problems in boys."

> "The current results suggested that phthalate exposure during pregnancy affected neurobehavior in children aged 0-12 years. Phthalates tended to have an impact on short-term cognitive development, which may be due to the fact that the assessment of the intelligence scale required special training and that follow-up is more difficult longer duration, resulting in a lack of literature on the long-term effects."

> "Prenatal exposure to DEHP, DBP, DEP, and BBzP had an adverse impact on cognitive development, psychomotor development, internalizing behavior, externalizing behavior, attention, gender-typical play behaviors, social behavior and visual spatial ability in children, but results are not always consistent. Some effects were gender-specific, especially the impact of phthalate exposure on neurobehavior in boys."

Radke et al. (2020) is a systematic review and meta-analysis of the relationships between phthalate exposure and neurodevelopment. The authors identified fourteen (14) studies on cognition, nine (9) studies on motor effects, twenty (20) studies on attention deficit hyperactivity disorder, three (3) studies of infant behavior, and seven (7) studies on social behavior. Any and all study design types were included. These authors did not use objective criteria to determine which studies were entered into their meta-analysis; instead, they relied upon "expert judgment." The authors concluded that there was not a clear pattern of association between prenatal phthalate exposure and neurodevelopment.

More Recent Studies of Phthalate Exposure and Neurodevelopment in Children

Because the most recent published reviews on phthalates and neurodevelopmental outcomes evaluated studies published through early 2019, it is important to examine the epidemiological studies published between 2019 and the present (i.e. March, 2020).

A systematic search of PubMed was undertaken on March 2, 2020 using the search terms "phthalates" and "neurodevelopment" to include English-language epidemiological studies. There were 51 publications identified, of which the following were considered relevant after full-text review: Dong et al. (2019), Hyland et al. (2019), Jankowska et al. (2019a), Jankowska et al. (2019b), Jones et al. (2018), Qian et al. (2019), and Tanner et al. (2020).

Brief descriptions of the conclusions of these studies follow.

Dong et al. (2019) conclude:

> "In conclusion, we found a significant negative association between lactational exposure to phthalates and ASQ-3 domains."

Hyland et al. (2019, p. 10) conclude:

> "Overall, we found mostly null associations of prenatal phthalate exposure with child and adolescent neurodevelopment outcomes in the CHAMACOS cohort, although we observed some suggestive associations of prenatal LMW phthalate biomarker concentrations with more internalizing and externalizing behaviors, particularly from self-reported and performance-based assessments. These findings add to a growing literature addressing the potential developmental neurotoxicity of phthalate exposure."

Jankowska et al. (2019, p. 6) conclude:

> "This study indicates that phthalate exposures in children from Poland might have an adverse effect on behavioral and cognitive development at early school age. Specifically, MMP and DnBP have been found to be related to poorer behavior scores from the SDQ test, as well as lower intellectual abilities from the IDS test."

Jankowska et al. (2019b, p. 7) conclude:

> "The present study confirms the influence of phthalate exposure, especially in the early postnatal period, on the development of the children. The consequences of phthalates'

> exposure depend on the characteristics of the environmental exposure itself, as well as on social, economic and demographic factors."

Jones et al. (2018, p. 4) conclude:

> "In summary, the expressive language scores in the Bayley's assessment of infant development for infants at 24 months of age are associated with levels of several metabolites in the maternal hair metabolome. These associations were not seen for receptive language, cognitive, or motor scores. As in previous studies, higher maternal phthalate exposure was associated with poorer outcomes. This finding highlights an early modifiable risk factor for compromised infant neurodevelopment that warrants further investigation."

Qian et al. (2019, p. 9) conclude:

> "We found a negative association between prenatal exposure to MnBP and psychomotor development in 2-year-old children and potential sex-specific associations with HMW phthalate levels among boys and girls."

Tanner et al. (2020, p. 9) conclude:

> "In conclusion we found that in a population-based pregnancy cohort, early prenatal exposure to a mixture of suspected EDCs was related to lower levels of cognitive functioning at age seven, particularly among boys."

Summary of Studies Published After Recent Reviews: Exposure to Phthalates and Neurodevelopment in Children

These studies add to the consistency of studies examining the relationship between exposure to phthalates during pregnancy and several neurodevelopmental outcomes.

Recent Reviews and More Recent Studies of Exposure to Polybrominated Diphenyl Ethers (PBDEs) and Neurodevelopmental Outcomes

Reviews of PBDE Exposure and Neurodevelopmental Outcomes: Lam et al. (2017), Gibson et al. (2018), and Vuong et al. (2018).

Lam et al. (2017) is a systematic review and meta-analysis focused solely on exposure to PBDE and neurodevelopmental outcomes (including but not limited to IQ and ADHD). The authors searched for epidemiological studies published through September 17, 2016 examining these relationships. Searching various databases (including PubMed, ISI Web of Science, Biosis Previews, EMBASE, Google Scholar, and TOXLINE), the authors identified English-language cohort, case-control, and cross-sectional studies; they excluded studies without a comparator, studies that did not have a quantitative measure of PBDE exposure, was a review, or did not examine neurodevelopmental outcomes. Only the most recent publication of any study was included; earlier published analyses of the same study were excluded. The authors used a modified version of the Cochrane Collaboration's "Risk of Bias" tool and the Agency for Healthcare Research and Quality's (AHRQ) domains to evaluate study quality. Four (4)

studies were identified and included in the meta-analysis regarding developmental exposure to PBDEs and IQ in children. The authors concluded (Lam et al. 2017, p. 17):

> "We found an association of PBDEs with decrements on IQ [3.7-point reduction in IQ per 10-fold increase (in other words, times 10) with a 95% confidence interval (-5.56 to -0.83 with an I^2 = 0%)] in PBDE exposure and concluded that there was 'sufficient' evidence supporting an association between developmental PBDE exposure and IQ reduction. Our findings suggest that preventing exposure to PBDEs could help prevent loss of human intelligence and, potentially, prevent other neurodevelopmental disorders in children."

See also Lam et al. (2017, Figure 3, p. 14).

In addition, the authors concluded that the evidence was of "moderate" quality for ADHD with "limited evidence" for an association with PBDEs, "based on the heterogeneity of association estimates reported by a small number of studies and the fact that chance, bias, and confounding could not be ruled out with reasonable confidence" (Lam et al. 2017, p. 1).

Gibson et al. (2018) is a narrative review of the associations between exposure to PBDEs and child cognitive, behavioral, and motor developmental outcomes. Sixteen (16) studies were identified that measured PBDE in maternal blood during pregnancy or in cord blood at delivery and used validated motor, cognitive and/or behavioral testing during childhood. The authors concluded that "the majority of studies support an adverse association between PBDEs and neurodevelopment" (Gibson et al. 2018, p. 1).

Vuong et al. (2018) is a narrative review of the association between exposure to PBDEs and child behavior. The authors identified 19 epidemiologic studies, 16 of which were prospective cohort in design. In these studies, PBDE exposure was measured at different development milestones, including gestation (maternal and cord sera), infancy (breastmilk), and childhood (serum). The outcome variables—behavioral batteries—were assessed in children ranging from 1 to 12 years of age. Outcomes included attention, executive function, and behavioral problems, such as externalizing, internalizing, adaptive and social behaviors, as well as autism spectrum disorder. The authors concluded the following (Vuong et al. 2018, p. 101):

> "The preponderance of the evidence from the available epidemiologic studies indicates that PBDEs affect behavioral development in children. PBDE exposure during gestational development is associated with impairments in executive function and poorer sustained attentional control in children."

<u>More Recent Studies of PBDE Exposure and Neurodevelopment in Children</u>

Because the most recent published reviews on PBDEs and neurodevelopmental outcomes evaluated studies published through late 2016, it is important to examine the epidemiological studies published between 2016 and the present (i.e. March, 2020).

A systematic search of PubMed was undertaken on March 16, 2020 using the search terms "PBDE" and "polybrominated diphenyl ethers" and "epidemiology" and "neurodevelopment" to include English-language epidemiological studies. There were 24 publications identified, of which the following were

considered relevant after full-text review: Braun et al. (2017), Chen et al. (2014), de Water et al. (2019), Ji et al. (2019), Lenters et al. (2019), Vuong et al. (2016), Vuong et al. (2017), Vuong et al. (2018), and Zhang et al. (2017). An additional study by Liang et al. (2019) was found in a separate search.

Brief descriptions of the conclusions of these studies follow.

Braun et al. (2017, p. 192) conclude:

> "We used repeated measures of neurobehavior in a cohort of typically developing children to shown how two prevalent neurotoxicant exposures may impact both the absolute level and trajectory of child behavior and cognition. Prenatal PBDE exposure was associated with persistent increases in externalizing behaviors from ages 2-8 years, declines in mental development from ages 1-3 years, and persistent decrements in IQ from ages 5-8 years."

Chen et al. (2014, p. 861) conclude:

> "In summary, the results of this study confirmed those of previous studies, which showed that prenatal exposure to PBDEs is significantly associated with reduced FSIQ in children. We found that prenatal exposure to PBDEs is also significantly associated with increases in scores of externalizing problems and hyperactive behaviors in children."

De Water et al. (2019, p. 1018) conclude:

> "Higher prenatal flame retardant concentrations were associated with more parent-reported executive functioning problems in children."

Ji et al. (2019, p. 5) conclude:

> "In the present longitudinal follow-up study, we found new human evidence of children's neurobehavioral problems following exposure to low PBDE levels, including somatic complaints, withdrawn, sleep problems, and internalizing problems in girls, and somatic complaints and attention problems in boys."

Lenters et al. (2019, p. 39) conclude:

> "We did not find clear evidence that PBDEs were associated with ADHD. We found suggestive evidence that BDE-47 was associated with an increased risk of ADHD and BDE-153 with a decreased risk."

The authors note that a potential explanation for their conflicting results is the fact that the PBDE levels in this Norwegian study population were much lower than in the U.S. populations where PBDE has been consistently associated with behavioral disorders.

Liang et al. (2019, p. 337) conclude:

> "Our study suggests that childhood serum PBDE concentrations, particularly BDE-153, may be inversely associated with children's reading skills."

Vuong et al. (2016, p. 2) conclude:

> "A 10-fold increase in BDE-153 was associated with poor behavior regulation (β = 3.23, 95% CI: 0.60, 5.86). Higher odds of having a score ≥ 60 in behavior regulation (OR = 3.92, 95% CI: 1.76, 8.73) or global executive functioning (OR = 2.34, 95% CI: 1.05, 5.23) was observed with increased BDE-153."

Vuong et al. (2017, p. 11) conclude:

> "Findings from this study are mixed as a pattern of impaired visual spatial learning was noted with increased concentrations of early childhood BDE-153, while improvements in visual spatial learning was observed with increased childhood concentrations of the Penta-BDE congeners BDE-47, −99, and −100 as well as BDE-28. Improved visual spatial memory retention was also observed with increased prenatal concentrations of BDE-28, −47, −99, and −100 as children spent more time and distance traveling in the correct quadrant. Results are suggestive of a sexually dimorphic relationship between PBDEs and visual spatial abilities, with male children performing more poorly on the Virtual Morris Water Maze. While the observed associations may have small effects at the individual level, at the population level the magnitude would be larger and thus be more meaningful."

Vuong et al. (2018, p. 1-2) conclude:

> "Null associations were observed between early childhood PBDEs and executive function. However, we observed significant adverse associations between a 10-fold increase in concurrent concentrations of BDE-28 (β=4.6, 95% CI 0.5, 8.7) and BDE-153 (β=4.8, 95% CI 0.8, 8.8) with behavioral regulation. In addition, PBDEs at 8 years were significantly associated with poorer emotional and impulse control. No associations were noted between childhood PBDEs and metacognition or global executive function. However, child sex significantly modified the associations, with significantly poorer executive function among males with higher concurrent BDE-153, and null associations in females. Our study findings suggest that concurrent PBDE exposures during childhood may be associated with poorer executive function, specifically behavior regulation. Males may also be more sensitive to adverse associations of concurrent PBDEs on executive function."

Zhang et al. (2017, p. 751) conclude:

> "In conclusion, prenatal PBDE concentration was inversely associated with children's reading abilities and FSIQ at age 8 years and positively associated with externalizing behavior problems at age 8 years."

Summary of Studies Published After Recent Reviews: Exposure to PBDEs and Neurodevelopment in Children

These studies add to the consistency of studies examining the relationship between exposure to PBDE during pregnancy and several neurodevelopmental outcomes.

Recent Reviews and More Recent Studies of Mercury and Neurodevelopmental Outcomes

Reviews of Mercury Exposure (including Methylmercury) and Neurodevelopmental Outcomes: Castoldi et al. (2008), Jurewicz et al. (2013), Grandjean and Landrigan (2014), Lanphear (2015), Silbergeld (2016), and Dorea (2019).

Castoldi et al. (2008) describe the results of studies of high and low dose Hg exposure and the profound effects Hg has on neurodevelopmental outcomes. Methyl mercury (MeHg) has been known to affect neurodevelopment at least since the food poisoning epidemics in Japan in the 1950's and 1960's and in Iraq in the 1970's. The Japanese epidemic involved consumption of Hg-contaminated fish from Minimata Bay. The Iraqi epidemic involved consumption of contaminated seed-grain. In both situations, major neurodevelopmental toxicity was observed in the offspring of mothers who ate fish (Japan) or grain (Iraq) during pregnancy. The adverse outcomes included microcephaly, cerebral palsy, blindness, deafness, dysarthria, abnormal reflexes and gross impairment of motor and intellectual development. Cognitive impairments did not improve over time. Lower levels of exposure to MeHG involved the fish-eating populations of the Seychelles in the Indian Ocean and the Faroe Islands between Iceland and Norway.

While the neurodevelopmental effects of low-dose exposure to MeHg are not as profound as those observed in the Japanese and Iraqi poisoning epidemics, the scientific community maintains a clear consensus that exposure to MeHg (and Hg) is harmful to the nervous systems of the developing fetus and children, as noted by the following:

> "Mercury is neurotoxic in any occurring chemical form, especially during neurodevelopment" (Dorea, 2019, p. 1292).
>
> "One of the most potent DNT (i.e. developmental neurotoxins) is methylmercury, a form of the metal produced in water and soils by bacteria" (Silbergeld, 2016, p. 7).
>
> "…mercury is an established risk factor for cognitive deficits" (Lanphear, 2015, p. 214).
>
> "Developmental neurotoxicity due to methylmercury occurs at much lower exposures than the exposures that affect adult brain function. Deficits at 7 years of age that were linked to low-level prenatal exposures to methylmercury were still detectable at the age of 14 years" (Grandjean and Landrigan, 2014, p. 331).
>
> "Methylmercury is a highly toxic substance; a number of adverse health effects associated with exposure to it have been identified in humans and in animal studies. Most extensive are the data on neurotoxicity, particularly in developing organisms" (US EPA, IRIS Website, accessed 4.2020).

More Recent Studies of Mercury Exposure (including Methylmercury) and Neurodevelopmental Outcomes:

Recent Reviews and More Recent Studies of PCBs (Polychlorinated Biphenyls) and Neurodevelopmental Outcomes

Reviews of PCBs and Neurodevelopmental Outcomes: Polanska et al. (2013), Schug et al. (2015), Vrijheid (2016), and US EPA (2020).

Polanska et al. (2013) reviewed evidence on the relationships between exposure to PCBs, organochlorine pesticides, organophosphate pesticides, and heavy metals and attention deficit/hyperactivity disorder in children, focusing on the epidemiological studies. The authors identified studies from searches on EMBASE, PubMed and EBSCO databases. The searches were limited to English-language epidemiological studies published since 2000. Thirty-one (31) studies were identified that examined these relationships. **The authors conclude that "the outcome of the presented studies has proven an association between exposure to organochlorine pesticides and PCBs and ADHD-like behaviors: alertness, quality of alert responsiveness, cost of attention, and other potential attention-associated measures including self-quieting and motor maturity"** (Polanska et al. 2013, p. 23)…(emphasis added).

Schug et al. (2015) is a narrative review examining the links between endocrine disruptors—EDCs—(including but not limited to PCBs) and neurodevelopment. These authors conclude that "**some of the most compelling evidence for the link between EDC exposure and impaired cognitive function involves PCBs,** chemicals that were widely used as coolants, plasticizers, and flame retardants, among many other uses, until their production was halted in the United States in 1979" (Schug et al. 2015, p. 1943)…(emphasis added).

Vrijheid et al. (2016) is a review of environmental pollutants and child health, including but not limited to PCBs. The authors searched the PubMed database using keywords, including "polychlorinated biphenyls" and "organochlorine compounds" and "neurodevelopment" and "cognition" among others. The authors conclude that "the evidence for associations between high exposures to prenatal PCBs (i.e. related to contamination incidents) and cognitive impairment is well established." In addition, the authors also concluded that "prenatal or early postnatal exposure to PCBs was associated with adverse neurodevelopmental outcomes and behavior in most of around 20 recent studies…"

According to the **US EPA** (**2020,** website www.epa.gov):

> "Proper development of the nervous system is critical for early learning and can have potentially significant implications for the health of individuals throughout their lives. Effects of PCBs on nervous system development have been studied in monkeys and a variety of other animal species. Newborn monkeys exposed to PCBs showed persistent and significant deficits in neurological development, including visual recognition, short-term memory and learning. Some of these studies were conducted using the types of PCBs commonly found in human breast milk."
>
> "Studies in humans have suggested effects similar to those observed in monkeys exposed to PCBs, including learning deficits and changes in activity associated with exposures to PCBs. The similarity in effects observed in humans and animals provide additional support for the potential neurobehavioral effects of PCBs."

More Recent Studies of PCB Exposure and Neurodevelopment in Children

Because the most recent published reviews on PCBs and neurodevelopmental outcomes evaluated studies published through 2014, it is important to examine the epidemiological studies published between 2014 and the present (i.e. 2020).

A systematic search of PubMed was undertaken on May 6, 2020 using the search terms "PCBs" and "polychlorinated biphenyls" and "children" and "neurodevelopment" to include English-language epidemiological studies. There were 68 publications identified, of which the following were considered relevant after full-text review: Caspersen et al. (2016), Kim et al. (2018), Kyriklaki et al. (2016), Lenters et al. (2019), Lynch et al. (2012), Nakajima et al. (2017), Ruel et al. (2019), Tsai et al. (2017), Wang et al. (2015), and Zhang et al. (2017).

Brief descriptions of the conclusions of these studies follow.

Caspersen et al. (2016, p. 649) conclude:

> "…our findings indicated that maternal dietary exposure to PCB-153 or dl-compounds during pregnancy was significantly associated with poorer expressive language skills in preschool girls, although the sex-specific associations were not significantly different."

Kim et al. (2018, p. 378) conclude:

> "Taken together, maternal exposure to several EDCs, such as PCBs and DEHP was associated with adverse neurodevelopmental performances among the children aged 1-2 years."

Kyriklaki et al. (2016, p. 210) conclude:

> "Overall, in the present study we found that prenatal exposure to HCB and PCBs was associated with reduced offspring cognitive development at preschool age."

Lynch et al. (2012, p. 455) conclude:

> "This study suggests that postnatal exposure to PCB 153 in breast milk, even at relatively low levels, may be associated with decrements in motor development among children after controlling for prenatal exposure and other relevant covariates."

Nakajima et al. (2017) studied dioxin-like compounds—"DLCs"—(including but limited to PCBs) and neurodevelopment in Japanese children. They conclude (Nakajima et al. 2017, p. 222):

> "These findings indicate that adverse neurodevelopmental effects of prenatal background-level exposure to DLCs may be stronger in male children."

Ruel et al. (2019, p. 6) conclude:

> "Higher prenatal levels of PCB-153 were associated with a delayed MDI score at 18 months."

Tsai et al. (2017, p. 396) conclude:

> "…PCBs have been shown to affect children's neurodevelopment…"

Wang et al. (2015, p. 407) conclude:

> "Postnatal exposure to PCBs was found to have potential adverse effects on the behavior and neurodevelopment of preschool-aged children…"

Zhang et al. (2017, p. 751) conclude:

> "Prenatal PCB concentration was not significantly associated with child's reading abilities, FSIQ, and externalizing behavior problems."

Summary of Studies Published After Recent Reviews: Exposure to PCBs and Neurodevelopment in Children

These studies add to the consistency of studies examining the relationship between exposure to PCBs and several neurodevelopmental outcomes.

Recent Reviews and More Recent Studies of Preterm Birth and Neurodevelopmental Outcomes

Reviews of Preterm Birth and Neurodevelopmental Outcomes: Linsell et al. (2015), Allotey et al. (2018), Burnett et al. (2018), Twilhaar et al. (2018), and Ylijoki et al. (2019).

Linsell et al. (2015) is a systematic review of risk factors for poor neurodevelopmental outcomes in very preterm (VPT) (≤ 32 weeks) or very low birth weight (VLBW) (≤ 1250 g) children survivors. Outcomes included were cognition, motor function, behavior, hearing and vision. Measures of interest included language skills, executive function, and academic achievement as well as global IQ. Because the factors that affect these outcomes are many the authors focus on studies with published multivariable outcome prediction models that aim to identify the combination of factors strongly associated with cognitive impairment in early infancy and later childhood. The authors identified relevant studies published between January 1, 1990 and June 1, 2014 reporting multivariable risk prediction models for a neurodevelopmental outcome assessed after age 18 months in VPT or VLBW children. Seventy-eight (78) articles representing results from forty-eight (48) cohorts were included in the review. Study quality was assessed using the "Quality in Prognosis Studies" tool and results were presented in accord with PRISMA guidelines. The authors found that male sex, nonwhite race/ethnicity, lower level of parental education, and lower birth weight were predictive of global cognitive impairment in children younger than 5 years. In older children, the primary risk factor was parental education.

Allotey et al. (2018) is a systematic review and meta-analysis of cognitive, motor, behavioral, and academic performance of children born preterm. The authors identified studies from PubMed and EMBASE published between January 1980 and December 2016; seventy-four (74) studies representing

64, 061 children were included in the review and meta-analysis. The general conclusions stated by the authors were that "there is strong relationship between gestational age at delivery and cognitive abilities, affecting moderately and late preterm infants. Furthermore, the neurodevelopmental deficits in preterm children persist beyond primary school age for all domains" (Allotey et al. 2017, p. 23). Specifically, the authors report that preterm children have lower cognitive scores for FSIQ (SMD: -0.70; 95% CI: -0.73 to -0.66), PIQ (SMD: -0.67; 95% CI: -0.73 to -0.60), VIQ (SMD: -0.53; 95% CI: -0.60 to -0.47). In addition, lower scores for preterm children in motor skills, behavior, reading, mathematics, and spelling were observed at primary school age. Gestational age at birth accounted for 38-48% of the variance.

In sum, "prematurity of any degree affects the cognitive performance of children born preterm" (Allotey et al. 2017, p. 16).

Burnett et al. (2018) is a review of the biological and social factors that can influence neurodevelopmental outcomes in preterm infants. Preterm refers to children born at <37 weeks gestation. The authors identified studies examining a variety of neurodevelopmental outcomes that commonly affect preterm infants, including neurosensory outcomes (cerebral palsy, impaired vision, impaired hearing, developmental delay in younger children and intellectual impairment in older children. They write that "cognitive deficits are the most commonly identified difficulties after preterm birth" (Burnett et al. 2018, p. 488). The factors that are responsible for these deficits include biological and social influences. As one example, the authors describe how in preterm infants, features of the Neonatal Intensive Care Unit (NICU), access to early developmental intervention, parental education and parental socioeconomic status together influence later cognitive functioning in preterm infants. The authors conclude (Burnett et al. 2018, p. 485):

> "Preterm birth (< 37 weeks of gestation) is a substantial risk factor for poor neurodevelopmental outcomes in areas such as neurosensory, cognitive, and behavioral functioning."

Twilhaar et al. (2018) is a meta-analysis of studies reporting on the association between preterm birth and academic performance. PubMed, Web of Science, and PsychINFO databases were searched to identify studies published since 1990 (when antenatal steroids and surfactant were available for preterm infants). Seventeen (17) studies were included in the meta-analysis and results were reported according to the PRISMA guidelines. Preterm children scored 0.71 standard deviations below their full-term peers on arithmetic ($p < 0.001$), 0.44 and 0.52 standard deviations below their full-term peers on reading and spelling ($p < 0.001$) and were 2.85 times more likely to receive special educational assistance ($p < 0.001$).

Ylijoki et al. (2019) is a systematic review focusing on the impact of several risk factors that adversely affect developmental outcomes in preterm infants, including clinical or histological chorioamnionitis, abnormal placental fetal and placental blood flow, and prenatal smoking exposure. Fifty-four (54) studies were identified published through September 12, 2018. Of these, the factor that showed adverse developmental outcomes in prenatal infants was maternal smoking.

Recent Reviews and More Recent Studies of Fluoride and Neurodevelopment in Children

Reviews of Fluoride and Neurodevelopment (Intelligence) in Children: Guth et al. (2020), Spittle (2020) and Duan et al. (2018).

A review (Guth et al. 2020), a commentary on that review (Spittle, 2020) and a meta-analysis (Duan et al. 2018) have recently been published examining the relationship between exposure to fluoride—typically in drinking water—and intellectual deficits (assessed by IQ) in children.

Guth et al. (2020) is a narrative review of 23 epidemiological studies published between 2012 and 2019 on fluoride and neurodevelopmental outcomes. Twenty studies were cross-sectional in design, one longitudinal and one prospective. IQ was the outcome measured in all but one study. The authors conclude that most studies were of low quality (e.g. insufficient control of confounding factors, no individual level exposure assessment) and inadequately designed to prove or disprove hypotheses (cross-sectional). Only two studies had what the authors consider to be suitable (i.e. prospective in design) were inconsistent in their results and controlled for some but not all relevant confounding factors. The authors note that 21 of 23 studies published since 2012 did reveal an association between exposure to fluoride and neurodevelopmental outcomes. The authors conclude that their review "does not support the presumption that fluoride should be assessed as a human developmental neurotoxicant at the current exposure levels in Europe.

Spittle (2020) is an editorial that discusses the differences between the conclusions of the Guth et al. (2020) review described above and an abstract published by Grandjean that purports to be an updated review of an earlier review by Grandjean and Landrigan (2014). Access to the full report by Grandjean is not available. Spittle (2020) nevertheless concludes that fluoride is a developmental neurotoxicant.

Duan et al. (2018) is a dose-response meta-analysis of twenty-six studies published up through 2016 on fluoride exposure and neurodevelopmental outcomes, specifically IQ. The authors conclude that "high levels of fluoride exposure significantly affected the development of intelligence in children" (Duan et al. 2018, p. 92) with a positive dose-response relationship.

More Recent Studies of Fluoride and Neurodevelopment in Children

Because the most recent published review on fluoride and neurodevelopmental outcomes evaluated studies published through 2019, it is important to examine the epidemiological studies published between 2019 and the present (i.e. October, 2020) that were not included in the Guth et al. (2020) review or the Spittle (2020) editorial.

A systematic search of PubMed was undertaken on October 14, 2020 using the search terms "fluoride" and "neurodevelopment" and "children" to include English-language epidemiological studies. There were 88 publications identified, of which the following were considered relevant after full-text review and had not appeared in the reviews described immediately above: Saeed et al. (2020), Lou et al. (2020), Till et al. (2020), and Wang et al. (2020).

Brief descriptions of the conclusions of these studies follow.

Saeed et al. (2020) conclude that their findings indicated that fluoride accounted for lower intelligent quotient.

Till et al. (2020, p. 1) conclude that "exposure to increasing levels of fluoride in tap water was associated with diminished non-verbal intellectual abilities."

Wang et al. (2020) observed statistically significant associations between fluoride exposure and diminished intelligence.

Summary of Studies Published After Recent Reviews: Fluoride and Neurodevelopment in Children

These studies add to the consistency of studies examining the relationship between exposure to fluoride and diminished intelligence.

OTHER FACTORS ASSOCIATED WITH OR CAUSING NEURODEVELOPMENTAL OUTCOMES

Recent Reviews and More Recent Studies of Prenatal Cocaine Exposure and Neurodevelopmental Outcomes

Reviews of the Effects of Prenatal Cocaine and Methamphetamine Exposure to the Developing Child

Smith and Santos (2016) is a narrative review of the effects of prenatal cocaine and methamphetamine exposure on the developing child. Both drugs are associated with behavioral rather than cognitive disorders. The authors write (Smith and Santos, 2016, p. 143-4):

> "Behaviorally, cocaine exposure is associated with increased symptoms of attention-deficit/hyperactivity disorder (ADHD)."
>
> "Further, cocaine exposure is associated with poor auditory attention skills and lower short term verbal memory scores."
>
> "Similar to cocaine, children exposed to methamphetamine do not have any deficits in IQ or language development. Behaviorally, methamphetamine exposure is associated with increased anxious/depressed problems and emotional reactivity, increased externalizing behaviors and ADHD symptoms and higher ADHD confidence index score, suggesting a greater future risk for developing ADHD."

Appendix B Scientific Reliability and Validity

In the event that I will be asked to examine the reliability and validity of the plaintiff's experts' opinions in this matter, I provide a description of two fundamental methodological concerns in the practice of science, with additional comments on bias, the most important challenge to the internal validity of scientific evidence.

Scientific Reliability

In science generally, and in the specific biomedical and public health sciences (e.g. epidemiology, toxicology, and biostatistics), reliability is defined as the extent to which a measurement technique or the results of a methodology can be replicated (Porta, 2008, p. 214). For example, a reliable method is one in which different individuals using the same method (under the same or similar conditions) report the same (or similar) results. Examples include blood tests, X-rays, exposure assessments, and the results of studies (e.g. the results of clinical trials or epidemiological studies). The results of meta-analyses and the results of other methods of research synthesis (e.g. the systematic narrative review) also fall under this definition of reliability (Weed, 2006).

Reliability, as defined here, tracks well with the definition of scientific objectivity, which emerges from (or can be traced to) the scientist's reliance upon methodology rather than upon subjective judgment. One of the best examples of the relative superiority of objective methodology over subjective judgment (or opinion) is the well-known hierarchy of scientific study designs utilized by the United States Preventive Services Task Force (USPSTF, 2008, p. 36) which puts randomized clinical trials and systematic reviews at the top of the hierarchy and expert opinion at the bottom (at or below the level of case reports) to mean that the objective methodologies that test causal hypotheses (and objectively assess bodies of evidence) trump the use of opinion and personal experience. The USPSTF hierarchy is as follows:

- Randomized Controlled Trials, well-conducted Systematic Reviews or Meta-Analyses of Homogeneous RCTs
- Well-designed Controlled Trials without Randomization
- Well-designed Cohort Studies
- Well-designed Case-Control Studies
- Multiple Time Series (a Type of Ecologic Study)
- Case Reports and Case Series, and
- Opinions of Respected Authorities (Experts) based upon "Experience."

Note that case reports, case series, and the opinions of respected authorities based upon "experience" are not tests of hypotheses. Indeed, a case series cannot test causal hypotheses for reasons discussed in earlier sections of this report.

Scientific Validity

In science generally, and in the specific biomedical sciences such as epidemiology, toxicology, and biostatistics, validity means the extent to which a measurement (e.g. a blood test, the diagnosis of disease, or the results of a study) measures what it purports to measure (Porta, 2008, p. 251-2). The main challenges to validity are systematic and random errors.

Validity, therefore, can refer to the extent to which a study (e.g. an epidemiologic study) is free from bias (also known as systematic error) and is free from random error (i.e. chance). Put another way, an internally valid study measures the risk of disease attributable to (i.e. caused by) the exposure, conditional on the absence of unknown confounding and other forms of bias. This version of validity—often called "internal validity"—is an assessment of the extent to which the reported results are free from errors (Aschengrau and Seage, 2003c).

Confounding is another source of systematic error that must be accounted for in an epidemiological analysis in order for the results of an analysis to be considered internally valid. Confounding can be due to unmeasured factors (associated with both exposure and outcome) or it can be due to incomplete measurement of factors, called "residual" confounding (Aschengrau and Seage, 2003d, p. 295).

Validity can also be used in a broader sense in science to describe the extent to which the results of a study—or more broadly, a body of evidence comprised of several studies from one or more disciplines—are generalizable. For a study or a body of evidence to be "externally valid" means that the results can be applied to populations other than the populations actually involved in the studies (Porta, 2008, p. 252; Aschengrau and Seage, 2003c). External validity, broadly conceived, is similar to the process of establishing general causation because a claim of general causation is equivalent to saying that the relationship at issue holds, with some obvious caveats (e.g. gender-specific illnesses such as cervical cancer), for any and all populations.

Appendix C

Literature Search and Included/Excluded Publications: Lead and Neurodevelopmental Outcomes

Appendix C contains two distinct yet interrelated efforts:

1. A systematic assessment of the peer-reviewed literature on the relationships between exposure to lead and neurodevelopmental outcomes, broadly interpreted to mean both cognitive and behavioral outcomes
2. A systematic assessment of the peer-reviewed literature on the relationship between exposure to lead and behavioral symptoms typically found in children diagnosed with attention-deficit hyperactivity disorder (ADHD) but also found in children without that diagnosis.

Appendix C Part One: Lead and Neurodevelopmental Outcomes (Cognition and Behavior)

The peer-reviewed literature available that evaluates the effects of lead [Pb] on health from a toxicological, epidemiological and clinical perspective is extensive. Due to the vast collection of studies available for review, the Agency for Toxic Substances and Disease Registry (ATSDR) Toxicological Profile for Lead (2007) was used as the starting point for this systematic review of lead exposure and cognitive outcomes. However, it must be noted that even the authors of the lead profile acknowledged the difficulty in reviewing every single reference available: “Due to the extent of the Pb database in humans, it is impossible to cite all, or even most, of the studies on health effects of Pb; thus, this profile does not attempt to provide a comprehensive review of all literature; instead, the profile summarizes the major lines of epidemiological evidence regarding health effects in humans.” (ATSDR, 2007 p. 11).

Two databases (PubMed and Web of Science) were used to conduct a literature search on human health effects due to lead exposure published as of January 1st, 2007 through December 31st, 2017. The focus of the literature search was to capture papers evaluating low blood lead levels and health effects. The PubMed search string included the following: ((("Lead"[Mesh]) AND ((((((((("cohort studies" [MeSH]) OR "longitudinal") OR "prospective") OR "nested case-control") OR "clinical trial") OR "meta-analysis" [publication type]) OR "systematic review") OR "case-control studies") OR "cross-sectional studies") OR "meta-analysis")) NOT ((("case reports" [publication type]) OR "editorial" [publication type]) OR "letter" [publication type)). This resulted in 807 research papers. A separate, but similar PubMed search was conducted to assess toxicology studies. This search found an additional subset of papers evaluating human health effects (n = 992) papers on human health and lead exposures from the PubMed database. Of the 1799 studies identified from PubMed, a total of 792 papers met the inclusion criteria.

Preliminary exclusion criteria included: article was published in a language other than English, non-human subject study, letter to the editor, conference abstract, paper included co-exposures to other chemicals, studies were case report or case series, there were no health endpoints reported (e.g. a study only reporting blood lead level distributions in a population), the study only evaluated exposures or conducted a risk assessment, the paper reported intervention techniques, the paper only discussed methodological, statistical or analytical techniques, or the studies assessed epigenetics. Topics and health categories of interest included papers specific to Flint, Michigan, and the following health endpoints: musculoskeletal, auditory, cancer, cardiovascular disease, dermatological, endocrine, fertility, hematological, hepatic, mortality, reproductive/developmental endpoints, neurological, ocular, renal, and respiratory.

This literature search was also conducted using comparable search terms with Web of Science. Over 25 different searches were conducted using the key words "lead" and "exposure" and the searches were filtered based on the topic areas described above, resulting in 2,378 potentially relevant articles. The systematic literature search described above was updated in October of 2019. Between the two databases (PubMed and Web of Science), an additional 444 articles were identified for screening. For this report, the focus is on the epidemiological literature that was directly related to low lead exposure (as reflected by mean blood lead levels for the population of interest ≤ 10µg/dL) and cognitive or neurobehavioral outcomes in children. Overall, there were 207 peer-reviewed papers that assessed lead exposure and these endpoints, however, 73 were excluded because the mean blood lead data distribution was higher than the designated cut point or proxy measures of exposure (e.g. soil lead levels or maternal blood lead levels) were used. In the end, 134 articles (60 cohort studies, 61 cross-sectional studies, and 13 case-controls studies) were reviewed.

As part of this review, each cohort study and case-control study were evaluated using the Newcastle-Ottawa Scale (NOS). The NOS was developed to aid researchers in assessing the quality of nonrandomized studies. The focus of the NOS is to ascertain the quality of three main categories: the selection of the study group, the comparability of groups, and the ascertainment of either the exposure of outcome of interest for case-control or cohort studies respectively (Wells et al. 2019).

Cohort Studies of Lead and Cognition in Infants and Toddlers

Cooney et al. (1989) conducted a study to investigate the possible effect of prenatal and postnatal lead exposure on neurobehavioral development in Australian children. The authors recruited 318 infants born in Sydney hospitals from 1982 to 1983 and followed them for 48 months. Prenatal lead exposure was measured by maternal and cord blood lead levels and postnatal lead exposure was measured by blood lead levels from ages six to 48 months. Neurobehavioral development was assessed through the McCarthy Scales of Children's Abilities at three and four years of age. Mean blood lead levels increased from birth (cord BLL = 8.1 ug/dL) to 18 months (BLL = 15.3 ug/dL) and then steadily decreased (48-month BLL = 10.1 ug/dL). Regression analyses were conducted to evaluate blood lead levels at various time points and developmental outcomes, adjusting for gestational age, verbal ability, education of mothers, education of fathers, occupational status of the father, and the total HOME score. The authors reported a correlation coefficient between prenatal cord blood levels at 12, 18, 24, and 30 months, and Mental McCarthy Scale (GCI) scores at 48 months. Additionally, pairs of measures were averaged to assess lead exposure over 12-month periods to remove any short-term fluctuations in blood lead levels, while still evaluating age-related trends. This analysis demonstrated increasing consistency with age, although there were no statistically significant negative correlations observed. The authors concluded that their study provided no evidence that low-level ambient lead exposure in the prenatal or early postnatal period is associated with mental or motor development at four years of age.

Rodrigues et al. (2016) performed a prospective birth cohort that was aimed to investigate the association between environmental exposure to lead, arsenic, and manganese and neurodevelopmental outcomes among 525 children in Bangladesh. Pregnant mothers were recruited from the Sirajdikhan and Pabna regions of Bangladesh between 2008 and 2011. Water samples from the wells used by the families were collected during the first trimester of pregnancy and again at follow-up visits at 1, 12, 20, and 40 months. Additionally, between 20 to 40 months of age, blood lead samples were taken from the children and neurodevelopmental outcomes were assessed using the Bayley Scales of Infant and Toddler Development, Third Edition (BSID-III). Neurodevelopment measures were cognition, receptive language, expressive language, fine motor, and gross motor. Analysis of children from this prospective birth cohort

was performed using linear regression and adjusted for relevant confounders, including the child's HOME score, maternal age, maternal education, maternal Raven score, exposure to environmental tobacco smoke, the child's sex, and child's hematocrit levels. The median age of children at the time of neurodevelopment assessment was 2.3 years; 49.7% of the children were males and 50.3% were females. Median blood lead concentrations were significantly higher in Sirajdikhan compared to Pabna (7.6 µg/dL versus below LOD µg/dL; p-value < 0.0001). The LOD for blood lead was 3.3 µg/dL and 67% of Pabna subjects had BLL less than this value. For comparison the 75th percentile BLL for children in Sirajdikhan and Pabna was 10.4 and 3.8 µg/dL, respectively. Increased blood lead was significantly associated with decreased cognitive scores in Sirajdikhan (β = –0.17, SE = 0.09, p = 0.05). Associations between lead and cognitive scores were not statistically different by clinic (p = 0.25). There was a statistically significant interaction observed between arsenic and lead in relation to cognitive scores (p = 0.003). The authors concluded that lead was associated with decreased cognitive BSID-III scores.

Ruiz-Castell et al. (2012) conducted a prospective cohort study aimed to evaluate lead exposures in 246 Bolivian children living near mining facilities and neurodevelopmental outcomes. Pregnant women were recruited between May 2007 and November 2009 from prenatal care units of two hospitals, located 800 and 1500 meters, respectively, from the San Jose's mine. Women were excluded if they had difficulty understanding the project and its objectives, were younger than 17 years, diagnosis of multiple pregnancies, not a resident of Oruro, pre-term children (<37 weeks), and had other children with behavioral problems. Prenatal blood lead levels were taken from the mother's at their hospital visit between the second and third trimester of pregnancy. The geometric mean of maternal blood lead levels was 1.76 ug/dL (95% CI: 1.68-1.84). Between 10.5 and 12.5 months of age, child neurodevelopment was assessed using the Bayley Scales of Infant Development (BSID). The authors used a multivariable linear mixed model to measure this association while adjusting for child gender, examiner, cesarean delivery, Bayley testing quality, and number of Bayley tests administered. After adjustment, there was a positive correlation between maternal lead levels and child mental (β =2.27, p=0.034) or psychomotor (β =2.74, p=0.012) development. The study authors concluded that no toxic effect of metals of neurodevelopmental outcomes was observed, which suggested that women from the mining region were not highly exposed.

Plusquellec et al. (2010) conducted a cohort study of 5-year old Intuit children from Arctic Quebec (N=110) in order to examine the association between prenatal and current exposures to Pb, PCBs, and Hg with behavioral outcomes. The current study was a follow-up of the Cord Blood Monitoring Program that took place between 1993 and 1996, and observations when children were 5 years old took place between January 2000 and October 2002. Cord blood samples were analyzed for the contaminants of interest, including lead, and used to represent prenatal exposures. In addition, blood samples were collected and analyzed at the follow-up visit that occurred when children were around four to six years of age. Mean cord blood lead level was 5.0 µg/dL (SD = 3.6) and mean child blood lead level was 5.4 µg/dL (SD = 5.0). Behavioral outcomes were assessed at the four to six years old follow-up visit and included a modified version of the IBRS from the BSID-II, as well as behavioral coding of videotaped behaviors. Any variable associated with an outcome at a p-value ≤ 0.20 following a Pearson's correlation analysis was included as a potential confounder in the first step of the multiple regression analysis. Final regression models were computed for each behavioral outcome and included prenatal and postnatal exposure to environmental contaminants, as well as the initial set of potential confounders that were then removed one by one if they were not significantly associated (p ≥ 0.10) with the outcome. Multiple regression models examined whether each of the environmental contaminants were independently associated with behavioral outcomes. Models that examined child blood lead adjusted for different

subsets of the following confounding variables: birthweight, child's sex, binge drinking during pregnancy, breastfeeding duration, caretaker RAVEN score, duration of pregnancy, parity, caretaker years of schooling, number of children at home, cord DHA/AA Docosahexaenoic acid/Arachidonic acid), Hollingshead index, child blood Hg, cord DHA, and child DHA/AA. After adjusting for confounders, it was found that cord blood lead concentrations were not associated with any behavioral outcomes. However, childhood blood lead concentrations measured at the time of testing were associated with higher impulsivity ($B = 0.20$, $p<0.05$) and higher irritability ($B = 0.20$, $p<0.05$) as measured by the IBRS, and a greater portion of time spent off task as measured by behavioral coding ($B = 0.21$, $p<0.05$). The authors concluded that the lack of agreement between prenatal lead exposure and behavioral outcomes, as well as the significant associations between lead exposure at the time of testing with increases in impulsivity, irritability, and inattention, corroborate previously reported results.

Liu et al. (2013) examined the relationships between blood lead concentrations and children's IQ and school performance in a cohort of 1,341 children (738 boys and 603 girls) from Jintan, China. Children aged 3-5 years attending four preschools were recruited between fall 2004 and spring 2005. Blood lead concentrations were measured between 3-5 years of age, IQ was assessed at age 6 using the Wechsler Preschool and Primary Scale of Intelligence – Revised (WPPSI-R), and school performance in Chinese, Math, and English was assessed between ages 8-10 years by standardized city tests. The mean blood lead concentration was 6.43 µg/dL (SD = 2.64). To examine the adjusted associations between blood lead concentration and IQ and school performance, generalized linear models (GLM) were performed, and controlled for various confounders, including child age at blood test, child gender, residence as defined as school location, blood iron level, parent education, parent occupation, and father's smoking history. Children with blood lead concentrations ≥10 µg/dL scored lower on IQ tests than children with blood lead concentrations <8 µg/dL, although the association was not significant after adjusting for potential confounders (PIQ: β = -1.90, 95% CI: -4.89, 1.09; VIQ: β = -1.77, 95% CI: -4.01, 0.46; FIQ: β = -1.45, 95% CI: -3.50, 0.67). However, when compared to children whose blood lead concentrations were <8.0 ug/dL, standardized test scores in all three subjects significantly declined with blood lead concentrations 8-<10 µg/dL (Chinese, β = -3.20, 95% CI: -5.78, -0.63; Math, β = -5.25, 95% CI: -8.14, -2.36; English, β = -4.33, 95% CI: -7.32, -1.34) and ≥10 µg/dL (Chinese, β = -4.02, 95% CI: -7.11, -0.93; Math, β = -5.27, 95% CI: -8.73, -1.81; English, β = -5.18, 95% CI: -8.76, -1.59). The study authors concluded that their findings support that blood lead concentrations, even <10 µg/dL, had negative impact on cognitive development.

Liu et al. (2014) evaluated the cognitive effects induced by low-level lead exposure during prenatal and postnatal periods in a cohort of 362 infants from the Pearl River Delta Region, Guangdong, China. Mother's delivering their babies in three centers were recruited from January 2009 to January 2010 and the following exclusion criteria was applied to the mothers: not a resident of the city, planning to leave the area within 5 years, daily consumption of alcohol, addiction to illegal drugs, continuous use of prescription drugs, diagnosis of multiple pregnancy, preeclampsia, renal or heart disease, gestational diabetes, and use of corticosteroids. Additionally, infants were required to meet the following criteria: absence of a medical condition considered to a risk factor for developmental difficulty, residence near the hospital in an area considered safe for home visitors, and maternal consent. Umbilical cord blood samples were collected after birth and infant blood samples were collected at 6, 12, 24, and 36 months. A total of 243 children (102 in the low lead group, 141 in the high lead group) had complete data, however there was reportedly no difference in the dropout rate between children with high (≥3.92 µg/dL) and low (≤1.89 µg/dL) cord blood lead levels. The Bayley Scales of Infant Development (BSID-II) were used to evaluate developmental functioning of the subjects during each of the four follow-up

points. The mean umbilical cord blood lead levels for the high and low lead group, respectively, were 5.63±0.32 µg/dL and 1.35±0.26 µg/dL. Multiple linear regression analyses were conducted to analyze the differences in developmental outcomes between the high and low lead groups. All models adjusted for father's occupational class, mother's education and IQ, and Child Home Nurture Environment (HNE) score. Different models for each outcome measured at each time point additionally adjusted for one to three of the following variables based on their statistical associations with the outcomes and cord blood lead levels: maternal hemoglobin, birth weight, gender, maternal age, ETS, mother's occupation, and household income. There were statistically significant inverse associations between umbilical cord blood lead levels and the mental development index (MDI) at 6 months (coefficient = -1.647, 95% CI: -2.094, -1.200), 12 months (coefficient = -1.458, 95% CI: -1.832, -1.084), 24 months (coefficient = -1.385, 95% CI: -1.683, -1.087), and 36 months (coefficient = -1.291, 95% CI: -1.550, -1.032), and the psychomotor development index (PDI) at 36 months (coefficient = -1.302, 95% CI: -1.572, -1.031). Additionally, it was reported that blood lead level taken at 24 months was significantly inversely associated with MDI at both 24 months (β = -1.403, p = 0.026) and 36 months (β = -1.298, p = 0.036), and that blood lead level taken at 36 months was significantly inversely associates with both MDI (β = -1.382, p = 0.032) and PDI (β = -1.201, p = 0.038) at 36 months. The authors concluded that their study demonstrated that prenatal and postnatal lead exposure as low as 5 ug/dL had an adverse effect on neurodevelopment.

Vigeh et al. (2014) conducted a cohort study in Tehran, Iran, to evaluate the effects of low levels of prenatal lead exposures in mothers and mental development in their children up to 3 years of age. A total of 364 mothers in their first trimester of a singleton pregnancy, and who were non-smokers, aged 16 to 35 years, and were free from chronic conditions (e.g. heart disease, hypertension, diabetes, cancer or renal failure) were recruited from three teaching hospitals in Tehran from October 2006 to March 2011. Maternal (n = 364) and umbilical cord (n = 224) blood lead levels were collected during pregnancy and delivery and mental development was assessed in 174 infants at 36 months old using the Harold Ireton Early Child Development Inventory. Mean maternal blood lead levels during the first, second, and third trimesters of pregnancy, respectively were 4.15, 3.44, and 3.78 ug/dL and the mean umbilical cord blood lead levels were 2.86 ug/dL. To examine whether levels of blood lead during pregnancy were independently associated with the risk of impaired mental development, logistic regression analysis was conducted, adjusted for hematocrit, maternal education level, body mass index, family income level, completed gestational age, birth weight, and birth order. After adjusting for covariates, there was a statistically significant relationship between increasing maternal blood lead levels in the first trimester of pregnancy with a low score of Early Child Development Inventory (OR = 1.74, 95% CI: 1.18-2.57). The authors concluded that the results of their study demonstrated that a relatively low level of prenatal lead exposure was associated with lower developmental scores in early childhood.

Schnaas et al. (2000) conducted a prospective cohort study aimed at evaluating the association between post-natal blood lead levels (BLLs) and the General Cognitive Index (CGI) score of the McCarthy Scale of Children's Abilities among children from Mexico City, Mexico. The authors additionally examined how this association varied over time. BLLs were measured in mothers throughout pregnancy and at the time of delivery, children's BLL were measured every 6 months from delivery to 54 months of age. Geometric mean BLLs (µg/dL) of the sample at 6-18 months, 24-36 months, and 42-54 months of age were 10.1 (range = 3.5-37.0), 9.7 (range = 3.0-42.7), and 8.4 (range = 2.5-44.8), respectively. GCI was assessed at 36, 42, 48, 54, and 60 months of age. Of the original cohort of 436 children, a total of 112 children with complete data for all the time points when GCI was assessed were included in the current analyses. It was noted that children who were not included in the present analyses had significantly higher geometric mean BLLs at 24-36 months and 42-54 months of age, as well as significantly lower GCI scores at 42, 48, and 54 months of age when compared to the children included in the present analyses. The

authors stated that these differences between the children included vs. those excluded from the analyses likely biased the results. Analyses were performed using the GLM for repeated measures and adjusting for sex, 5-min Apgar score, birth weight, birth order, educational level of mother, family SES, and maternal IQ. No significant effects were observed using prenatal or perinatal blood lead measures. There was no significant between-subjects effect of BLL at 6-18 months and GCI score at any age. It was reported that there was an inverse association between BLL at 24-36 months and GCI score at 48 (p=0.021) and 54 months (p=0.073), as well as an inverse association between BLL at 42-54 months and GCI score at 54 (p=0.040) and 60 (p=0.060) months. A significant within-subjects inverse relationship was observed between BLL at 6-18 months and GCI score up to 48 months, as well as between BLL at 24-36 months and GCI score up to 48 months. The authors concluded that the effect of postnatal BLL on GCI peaks around 1 to 3 years following exposure, and then attenuates over time. Further, they indicated that 4 to 5 years of age appeared to be critical period for the manifestation of prior postnatal lead exposure.

Tellez-Rojo et al. (2006) conducted a prospective cohort study that aimed to evaluate the exposure-response relationship between blood lead levels (BLLs) and neurodevelopment at 12 and 24 months of age among infants from Mexico City, Mexico. Children were members of either a cohort that recruited pregnant women from Mexico, City between January 1995 and June 1995, or between May 1997 and July 1999. Pregnant women were excluded from either cohort if they reported daily alcohol consumption, addiction to illegal drugs, use of corticosteroids, or continuous use of prescription drugs, as well as if they received diagnoses of multiple pregnancy, preeclampsia, renal or heart disease, gestational diabetes, or seizures requiring medical treatment. Cord BLL was measured as well as children's BLL at 12 and 24 months of age. Primary outcomes were the Mental Development Index (MDI) and Psychomotor Development Index (PDI) scores derived from the Bayley Scales of Infant Development II (BSID II) that was administered to children at 12 and 24 months of age. The current analyses included 294 children with BLLs < 10 µg/dL at both 12 and 24 months, gestation of 37 weeks or more, birth weight greater than 2000 g, and completed data for BLL, outcome metrics, and maternal IQ. Mixed effects regression models adjusting for birth weight, gender, child's age, cohort membership, and maternal IQ were used to estimate the associations between BLL at 12 months of age with MDI and PDI scores at 12 and 24 months of age, as well as the cross-sectional association between BLL at 24 months of age with MDI and PDI scores at 24 months of age. Natural log-transformed BLL values were used in the models. Average BLL among participants was 4.27 (SD = 2.14) µg/dL at 12 months and 4.28 (SD = 2.25) µg/dL at 24 months. The results indicated that BLL at 12 months was not associated with MDI at 12 months or 24 months, or PDI at 12 months; however, BLL at 12 months was shown to be associated with lower PDI at 24 months (B = -3.0, p = 0.01). Further, the results showed that there was a significant inverse association between BLL at 24 months and MDI at 24 months (B = -4.7, p < 0.01), as well as between BLL at 24 months and PDI at 24 months (B = -5.44, p < 0.01). Additional analyses showed that the magnitude of the decreases in both MDI at 24 months and PDI at 24 months was significantly greater for children with concurrent (24-month) BLLs < 10 µg/dL compared with concurrent (24-month) BLLs ≥ 10 µg/dL. The authors concluded that neurodevelopment in children is inversely related to blood lead levels, even in the range below 10 µg/dL.

Braun et al. (2012) examined the pattern of associations between serial blood lead concentrations and children's cognitive abilities at four years of age to assess whether blood lead concentrations at different times during early childhood were more strongly associated with deficits in cognitive abilities. A total of 1,035 mother-child pairs were recruited from four prospective birth cohorts between 1993 and 2006 in Mexico City, Mexico. All enrolled women were from low to moderate income households and women

were excluded if they planned to leave the area within five years; had a history of infertility, diabetes, or psychosis; consumed alcoholic beverages daily during pregnancy; were addicted to illegal drugs; were diagnosed as a high-risk pregnancy; and were pregnant with multiples. Blood lead was measured in children at 1, 2, 3, and 4 years of age, and child cognitive abilities were evaluated at 4 years of age using the general cognitive index (GCI) of the McCarthy Scales of Children's Abilities. The authors utilized multivariable linear regression to estimate the change in cognitive abilities at 4 years of age with a 10 µg/dL increase in childhood blood lead concentrations, and adjusted for maternal education, maternal IQ, marital status, child sex, breast feeding duration, and study cohort. The average blood lead concentration among children with complete follow-up at four years of age was slightly higher than those with missing data (geometric mean of 5.3 µg/dL versus 4.9 µg/dL; 95% CI: 0 to 13%). Median blood lead concentrations among children at 1, 2, 3, and 4 years of age were 4.2, 4.6, 5.5, and 5.9 µg/dL, respectively. Children were 48.4 months of age on average at follow-up. The blood lead concentrations at 2 years were most strongly associated with reduced GCI scores at 4 years (β = –3.8; 95% CI: –6.3 to –1.4). Some of the associations between blood lead and GCI varied in magnitude and direction through cohorts. The authors concluded higher blood lead concentrations among children 2 years of age were most predictive of decreased cognitive abilities in this cohort, after adjusting for confounders.

Claus et al. (2012) prospectively examined manganese-lead interactions in early childhood to evaluate whether co-exposures from these metals was associated with neurodevelopmental deficiencies. A total of 486 eligible children in Mexico City, Mexico, were enrolled in a longitudinal cohort study from 1997 to 2000 and followed from birth through 36 months of age. Children with very low birth weight (<1,500 g) and severely premature birth (<32 weeks gestation) were excluded from the study. Blood lead and manganese concentrations were measured in children at 12 and 24 months of age, and child neurodevelopment was assessed using the Bayley Scales of Infant Development-II (BSID-IIS) at 12, 18, 24, 30, and 36 months of age. The mean blood lead concentrations in the umbilical cord, 12 months, and 24 months, respectively, were 4.7, 5.1, and 5.0 µg/dL. The authors utilized linear mixed-effects models with repeated measures of Bayley scores to examine the manganese-lead interaction over time, and also fit adjusted regression models for each exposure time point with each Bayley assessment. Models adjusted for sex, hemoglobin, gestational age, maternal IQ, and maternal education. There were not significant interactions between MDI scores and lead concentrations at 12 months (β = -0.07, 95% CI: -0.39, 0.25) and 24 months (β = -0.08, 95% CI: -0.46, 0.30) or between PDI scores and lead concentrations at 12 months (β = -0.27, 95% CI: -0.56, 0.02) and 24 months (β = -0.18, 95% CI: -0.53, 0.17). The authors concluded that co-exposure to lead and manganese was associated with greater deficits in mental and psychomotor development than expected based on effects of exposure to only lead or manganese. There were no conclusions provided for pure lead exposures and developmental outcomes.

Parajuli et al. (2013) investigated the association between *in utero* toxic lead, arsenic, and zinc levels and neurodevelopmental indicators after birth in children from Chitwan Valley, Nepal. Of 119 eligible women recruited between September to October 2008, a total of 100 at-term (≥37 weeks of gestation) pregnant women who were between 18 and 40 years of age, lived in Chitwan for at least two years, singleton, and were free from diabetes, hypertension, at preeclampsia agreed to participate in the study. Cord blood samples were taken at birth from the infants to measure lead concentrations and neurodevelopment was assessed the day after birth by the Brazelton neonatal behavioural assessment scale (NBAS III). The NBAS III comprises seven clusters: habituation, orientation, motor system, state organization, state regulation, autonomic stability, and abnormal reflex. Multivariable models were conducted to examine the association between neurodevelopment indicated and cord lead concentrations, adjusting for mother's age, parity, mother's education level, log annual family income, mother's BMI, birth weight of babies, gestational age, and age of baby at NBAS assessment. Mean cord

blood lead level was 31.7 µg/dL and median cord blood lead level was 20.6 µg/dL. After adjusting for potential confounders, only the motor system cluster was significantly inversely associated with cord blood lead levels (coefficient = -2.15; 95% CI: -4.27, -0.03). The authors concluded that high levels of lead exposure during the prenatal period may induce retardation during *in utero* neurodevelopment.

Jedrychowski et al. (2008) conducted a cohort study to assess the neurocognitive status of infants at 6 months whose mothers were exposed to lead during pregnancy. Women who were aged 18 to 36 years with singleton pregnancies, non-smokers, and free from chronic diseases were recruited during visits at ambulatory prenatal clinics in Krakow, Poland, between January 2001 and March 2003. Lead levels in cord blood were assessed and visual recognition memory (VRM) was measured at 6 months using the Fagan Test of Infant Intelligence (FTII) in a total of 452 infants. The mean cord blood lead level was 1.42 µg/dL (95% CI: 1.35-1.48). The association between Fagan ranked scores and lead levels in cord blood were analyzed by linear regression analysis and Spearman rank correlation. Additionally, multiple logistic regression models were conducted, in which lead exposure was entered as a continuous variable and as a dichotomized variable, and adjusted for maternal education, parity, and gender of child. The VRM scores were inversely related to the infant's cord blood lead level (Spearman correlation coefficient = -0.16, p = 0.007). Furthermore, infants with a high lead level (>1.67 µg/dL) had a statistically significant increased risk (OR = 2.33; 95% CI: 1.32-4.11) for developmental delay when compared to those infants with low lead levels (≤1.67 µg/dL). The authors indicated that their study results suggested a subtle neurotoxic impact of low-level prenatal lead exposure occurred in infants from the Krakow inner city area.

Jedrychowski et al. (2009) evaluated the potential association between very-low-level prenatal lead exposure and mental development of Polish children at 12, 24, and 36 months of age. The cohort included 444 infants of mothers who were recruited in their first and second trimester of pregnancy during visits to ambulatory prenatal clinics in Krakow, Poland, between January 2001 and February 2004. Women between the ages 18 and 36 years with singleton pregnancies, non-smokers, and free from chronic diseases were eligible for enrollment. Cord blood samples were taken at birth and mental development was assessed using the Bayley Mental Development Index (MDI) at 12, 24, and 36 months of age. In order to assess the effect of maternal lead exposure during pregnancy and children's MDI scores, analyses were performed using generalized estimating equations (GEE) model, which adjusted for prenatal environmental tobacco smoke exposure, birth order, maternal education, and child gender. Subjects from the higher exposure group had significantly lower maternal education and MDI scores compared to the lower exposure group. The median lead level in cord blood was 1.23 µg/dL. The association between MDI scores and lead exposure were of borderline significance at 12 months of age (β = -5.42, 95% CI: -11.19, 0.35), and showed a significant inverse association between mental function and lead exposure among children at 24 months of age (β = -7.65, 95% CI: -14.68, -0.62) and 36 months of age (β= -6.72, 95% CI: -12.5, -0.89). The authors concluded that the results of their study indicated that the neurotoxic impact of very low levels of prenatal lead exposure (below 5 µg/dL) may occur in infants and very young children.

Jedrychowski et al. (2009) were interested in assessing the relationship between very low-level prenatal lead exposure and gender specific cognitive deficits in the first years of life. Participants were from a prospective cohort study that initially included 505 children born between January 2001 and February 2004 from women living in the inner city and outlying residential areas of Krakow, Poland. Women with pregnancy-related diabetes or hypertension, and current smokers were excluded from the cohort (Jedrychowski et al. 2003). Cord blood lead level was measured at birth and the Mental Development Index (MDI) from the Bayley Scales of Infant Development-II (BSID-II) was administered at 12, 24, and 36

months of age to evaluate cognitive deficits. The current analysis included a sample of 457 children with complete data. Linear multivariate regression and the generalized estimating equations longitudinal panel model were utilized to examine potential associations between prenatal lead exposure and each MDI score. Potential confounders adjusted for in the linear multiple regression models included maternal education, parity, gender, duration of breastfeeding, and prenatal and postnatal environmental tobacco smoke (ETS) exposure. Median cord blood lead level was 1.21 µg/dL and ranged from 0.44 to 4.60 µg/dL. Results indicated that boys had significantly lower MDI scores than girls at each of the follow-up time points. Maternal education, parity, and prenatal or postnatal exposure to ETS did not differ by gender. Results showed a significant gender by cord blood lead level interaction associated with MDI score at 36 months (β = –4.46; 95% CI: –8.28 to –0.63). Gender stratified multiple linear regression models adjusted for covariates showed a significant inverse association between prenatal lead exposure (cord blood lead level > 1.67 µg/dL) and cognitive function at 36 months of age among boys (β = –6.158; 95% CI: –9.946 to –2.1370), but not among girls (β = –0.738; 95% CI: –4.796 to 3.319). Additionally, the average deficit of cognitive functioning in the entire sample throughout the first three years of life was significantly associated with higher prenatal lead exposure (β = –3.00; 95% CI: –5.22 to –0.70). The authors concluded that there may be no threshold for lead toxicity in children, and that boys are more susceptible to lead-induced cognitive deficits than girls.

Al-Saleh et al. (2008) conducted a longitudinal study to evaluate the effect of prenatal and/or postnatal lead exposure on early cognitive development among selected group of children from birth to 2 years of age in Al-Kharj, Saudi Arabia. A total of 653 umbilical cord blood samples were obtained from healthy mothers aged 17-46 years between March and July 2004. Infants that were born with Down's syndrome, retinoblastoma, cleft palate or a problem likely to require hospitalization beyond 3 weeks of age, including respiratory distress syndrome and gestational age less than 34 weeks, were excluded from the study. The mean blood lead level was 2.21 ug/dL (SD: 1.691). Blood lead levels and cognitive development scores using the Bayley Scales for Infant Development were repeatedly measured in study subjects every six months from birth to two years of age, however, the study authors did not report the results of the cognitive development scores. Furthermore, multiple regression models were conducted to investigate risk factors affecting cord blood level levels in subjects with lead levels above the 75th percentile (2.475 ug/dL). When controlling for newborn's head circumferences, it was reported that there was a significant inverse association between cord blood and newborn's head circumferences (*B* = -0.158, *p* = 0.036). The study authors concluded that their study results indicated that in utero lead exposure, even at low exposure levels, may cause adverse effects on the growth and development in young children.

Canfield et al. (2003) conducted a prospective cohort study that aimed to examine the association between low-level exposure to lead and children's intellectual performance at three and five years of age. Study participants lived in Rochester, NY, were born between July 1994 and January 1995, and were recruit at 24-30 months of age from a previous dust-control trial that they were enrolled at 5 to 7 months of age. Children's blood lead levels were measured at 6, 12, 18, 24, 36, 48, and 60 months of age, and four exposure variables were constructed for use in the analyses including lifetime average, peak, concurrent, and average infancy blood lead concentrations. Mean blood lead concentration was lowest at 6 months (3.4 µg/dL), highest at 2 years (9.7 µg/dL), and decreased to 6.0 µg/dL at 5 years of age. Lifetime average blood lead concentration was 7.7 µg/dL at three years of age and 7.4 µg/dL at 5 years of age. 57% of children had peak blood lead concentration below 10 µg/dL at three years of age and 55.8% had peak blood lead concentrations below 10 µg/dL at five years of age. Children's intelligence was assessed at three and five years of age with the fourth edition Stanford-Binet Intelligence Scale, and the composite score was used to represent IQ. Mixed-model methods were used,

and all adjusted for the following covariates: child's sex, birthweight, iron status, mother's IQ, mother's years of education, race, tobacco use during pregnancy, yearly household income, and total HOME score. Additional linear models were estimated for the subgroup of children with peak lead concentrations below 10 µg/dL. After adjusting for covariates, significant inverse associations were found between IQ and all of the blood lead variables and there were no significant differences according to age (lifetime average BLL: β = -0.46, p = 0.004; peak BLL: β = -0.23, p = 0.01; concurrent BLL: β = -0.46, p = 0.002; average infancy BLL: β = -0.43, p = 0.02). When analyzing the subset of children with peak blood lead concentrations below 10 µg/dL, it was found that there was a significant inverse association between IQ and lifetime average, peak, and concurrent blood lead concentrations at both three and five years of age after adjusting for covariates. The analysis of this subgroup of children resulted in larger decreases in IQ score for each 1 µg/dL increase in blood lead concentration (β = -1.37, p = 0.03; β = -1.40, p = 0.005; β = -1.58, p = 0.001; for lifetime average, peak, and concurrent blood lead concentrations, respectively). The study authors concluded that even low blood lead concentrations (<10 ug/dL) were inversely associated with children's IQ scores at three and five years of age, and that associated declines in IQ are greater at these concentrations than higher concentrations.

Canfield et al. (2004) examined the association between pediatric lead exposure and neurocognitive functioning at 5.5 years of age from participants of a previous study (Lanphear et al. 1999). Participants were originally enrolled at 5 to 7 months of age in a previous study of lead dust control efficacy and were invited to participate in a 5-year neurobehavioral study when children reached 24 to 30 months of age. Blood lead levels (BLLs) were measured semiannually from 6 to 24 months of age and then measured at 36, 48, and 60 months of age, which were then used to calculate lifetime average BLL for analyses. Neurocognitive functioning was assessed at 5 and a half years of age using the Working Memory and Planning Battery of the Cambridge Neuropsychological Test Automated Battery (CANTAB), which consisted of six tasks: motor screening task (MOT), big-little circle task (BLC), spatial span task (SSP), spatial working memory task (SWM) and intra-dimensional-extradimensional shift task (IED). General linear, logistic, and mixed regression models were used to explore the associations between lifetime average BLL and cognitive outcomes. Numerous covariates considered for inclusion in the models: NICU admission, maternal IQ, HOME total score at 24 months, HOME cognitive simulation sub-score at age 6, breastfeeding duration, smoking during pregnancy, household income, number of residence changes, child's sex, birthweight, maternal ethnicity, marital status, maternal education, average crowding in the home, first prenatal visit, IED stages completed, SOC level completed, SSP span length, SWM problem type, and age at testing. Mean lifetime average BLL of the sample was 7.2 µg/dL with a range of 1.4 to 19.9 µg/dL. After adjustment for confounders, the following significant associations were observed with lifetime BLL: percent correct choices in BLC (β = -0.619, p<0.001), total nontarget errors made on SSP (β = 0.145, p<0.001), total errors made on SSP (β = 0.112, p = 0.016), total errors made on 6-box problems of SWM (β = 0.456, p = 0.016), stages completed on IED (β = -0.112, p = 0.025), completed ED shift on IED (β = -0.189, p = 0.023), total trials on IED (β = 1.40, p = 0.003), average moves to complete the SOC (β = 0.050, p= 0.009), and planning time for two-move problems on the SOC (β = 0.142, p < 0.001). The authors concluded that lead exposure was associated with impaired neuropsychological test performance among young children. They stated that their results suggest that lead exposures may be less damaging to sensorimotor functions compared with higher cognitive processes including focused attention, working memory, and other executive functions.

Jusko et al. (2008) conducted a prospective cohort study in order to examine the association between blood lead concentrations throughout early childhood and IQ at 6 years of age. Study participants lived

in Rochester, NY, were born between July 1994 and January 1995, and were recruited at 24-30 months of age from a previous trial of children that were first enrolled at 6 months old in a previous dust-control trial. Children's blood lead levels were measured at 6, 12, 18, 24, 36, 48, 60, and 72 months of age; however, prenatal maternal blood and umbilical cord blood specimens were not available. Intelligence was assessed with the Wechsler Preschool and Primary Scale of Intelligence, Revised (WPPSI-R) test that was administered at 6 years of age by an examiner that was unaware of the children's blood lead concentrations. Blood lead levels were used to construct four exposure variables that were used in analyses: lifetime average blood lead concentration, concurrent blood lead concentration (concentration measured on the same day of the WPPSI-R exam), infancy average blood lead concentration, and peak blood lead concentration. The outcomes of interest included Full-Scale IQ, Performance IQ, and Verbal IQ scores measured with the WPPSI-R. General linear models were used to estimate the association between each of the four blood lead exposure variables and the IQ scores at 6 years of age. Blood lead was modeled categorically and separated into categories of <5 µg/dL, 5.0-9.9 µg/dL, and ≥10 µg/dL; additional categories for peak blood lead concentration included 10.0-14.9 µg/dL and ≥15 µg/dL. 55% of children never had a blood lead concentration that was ≥10 µg/dL throughout the study period. All models adjusted for the following confounders: family income, child's sex, mother's highest level of education, race, prenatal smoking, birth weight, transferrin saturation, mother's IQ, and the HOME-SF score measured at 6 years of age. A generalized additive model with a locally weighted scatterplot smooth (LOESS) was used to estimate the dose-response relationship between peak blood lead levels and Full-Scale IQ. There was a significant inverse association between lifetime average blood lead concentration and Full-Scale IQ ($p = 0.006$) and Performance IQ ($p = 0.002$), between concurrent blood lead concentration and Full-Scale IQ ($p = 0.03$) and Performance IQ ($p = 0.004$), between infancy average blood lead concentration and Full-Scale IQ ($p = 0.05$) and Performance IQ ($p = 0.02$), and between peak blood lead concentration and Full-Scale IQ ($p = 0.03$) and Performance IQ ($p = 0.02$). There were no statistically significant associations between any of the blood lead exposure variables and verbal IQ score. Children with lifetime average blood lead concentrations between 5 to 9.9 µg/dL scored 4.9 points lower on Full-Scale IQ ($p = 0.03$) and 4.9 points lower on Performance IQ ($p = 0.03$) compared with children whose lifetime average blood lead concentrations were <5 µg/dL. Children with concurrent blood lead concentrations between 5 to 9.9 µg/dL scored 3.7 points lower on Full-Scale IQ ($p = 0.10$) and 5.5 points lower on Performance IQ ($p = 0.01$) compared with children whose concurrent blood lead concentrations were <5 µg/dL. Children with infancy average blood lead concentrations between 5 to 9.9 µg/dL scored 5.2 points lower on Full-Scale IQ ($p = 0.02$) and 5.4 points lower on Performance IQ ($p = 0.01$) compared with children whose infancy average blood lead concentrations were <5 µg/dL. Results of the non-linear analysis of peak blood lead concentration and Full-Scale IQ showed a significant inverse association that was apparent down to 2.1 µg/dL ($p = 0.003$). IQ score decreased by 1.2, 0.32, and 0.15 points per 1 µg/dL increase in peak blood lead over the ranges of 2.1-10 µg/dL, 10-20 µg/dL, and 20-30 µg/dL, respectively.

Desrochers et al. (2018) is the report of a Canadian study of 609 mother-child pairs from the Maternal-Infant Research on Environmental Chemical (MIREC) study. Women were recruited into the study during the years 2008 and 2011. Maternal and cord blood lead concentrations were obtained (1st trimester, 3rd trimester, and birth). Venous blood leads on the children in the study were obtained when the child was between 3 and 4 years of age; these levels were relatively low. The median blood lead concentrations for this cohort at 1st trimester, 3rd trimester, cord blood, and child blood were 0.60 µg/dL, 0.58 µg/dL, 0.79 µg/dL, and 0.67 µg/dL respectively. The Wechsler Preschool and Primary Scale of Intelligence was administered; that test is designed for children between 2 years, 6 months and 3 years. Potential confounders of the relationship between lead and IQ considered by these investigators were marital status, household income, education, parity, ethnicity, current smoking and alcohol

consumption during pregnancy, and exposure to mercury. Note that potential confounders not considered in these analyses were the following: exposure to PCBs, PBDEs, arsenic, manganese, and fluoride, adequate prenatal care, small for gestational age, exposure to chlorpyrifos and organophosphate pesticides, and exposure to phthalates to name most of the known risk factors for IQ changes in children. The authors reported no association between cord blood lead levels, child blood lead concentrations and IQ. See Desrochers et al. (2018, Table 3, p. 1239). As a secondary analysis and with no evidence to support it, the authors examined the extent to which gender moderated the relationship and found that the lead-Performance IQ relationship was not present in girls but present in boys.

The authors conclude (Desrochers et al. 2018, p. 1241):

> "Our study highlights that Canadian preschoolers from mainly middle-to upper-middle class families are exposed to particularly low levels of lead, and that their IQ was not associated with concurrent blood lead concentrations."

Cohort Studies of Lead and Cognition in School Age Children

Baghurst et al. (1992) were interested in the potential association between lifetime lead exposure and intelligence among children at seven years of age. Participants were recruited into a prospective birth cohort in the lead-smelting community of Port Pirie, Australia from 1979 to 1982, and a total of 494 children were evaluated at a follow-up between seven to eight years of age. It was reported that the children included in the present analyses had slightly more advantaged backgrounds compared to the children lost to follow-up. Blood lead levels (BLL) were measured before birth, from the umbilical cord at delivery, and when the child was aged 6, 15, and 24 months of age, and annually after 2 years of age. The revised Wechsler Intelligence Scale for Children (WISC-R) was administered to children between the ages of seven to eight years to assess IQ. Multiple regression models adjusted for sex, parents' education level, mother's age at birth, parents' smoking status, socioeconomic status, HOME scores, mother's IQ, birth weight, birth order, feeding style, breast-feeding duration, and whether the child's parents were living together. The mean lifetime averages of blood lead concentration at seven years of age calculated by quartile were 10.8 µg/dL, 15.7 µg/dL, 19.7 µg/dL, and 24.8 µg/dL. BLL was natural log-transformed for the statistical analyses. The exposure variable for BLL, log lifetime average blood lead concentration, was calculated as the area under each child's blood lead curve from birth until each age of interest (15 months, 2, 3, 4, and 7 years). After adjustment for confounders, it was shown that there were no significant associations between prenatal and cord BLL with children's IQ at age 7. The results of the adjusted models further showed that log lifetime average BLLs from birth to 15 months, 2, 3, and, 4 years of age were inversely significantly associated with verbal IQ and full-scale IQ ($p<0.05$). For an increase in blood lead concentration from 10 µg/dL to 30 µg/dL, the decrements in Full Scale IQ ranged from 4.4 to 5.3 points and decrements in Verbal IQ ranged from 5.5 to 6.4 points. It was noted that girls were more sensitive to lead exposure effects than boys, as indicated by the expected covariate-adjusted decrement in full-scale IQ being 7.8 points in girls and 2.6 points in boys for an increased blood lead concentration from 10 µg/dL to 30 µg/dL. The authors concluded low-level lead exposure during early childhood was inversely associated with neuropsychological development in the first seven years of life.

Schnaas et al. (2006) conducted a prospective cohort study in order to examine the association between prenatal and postnatal lead exposure with intellectual development in a group of children that were born in Mexico City between 1987 and 1992. A total of 321 children met the inclusion criteria of being

born with a gestational age of ≥36 weeks, 5-min Apgar score ≥6, birth weight >2000 g, no congenital anomalies, or the product of multiples births. Prenatal blood lead levels (BLLs) were measured at 12, 20, 28, and 36 weeks of pregnancy, cord BLL was measured at delivery, and child BLLs were measured every 6 months after birth to age 5 and then annually from ages 6 to 10. The geometric mean (range) during pregnancy, from 1-5 years, and from 6-10 years, respectively, was 8.0 µg/dL (1-33 µg/dL), 9.8 µg/dL (2.8-36.4 µg/dL), and 6.2 µg/dL (2.2-18.6 µg/dL). Child intelligence was measured by the Wechsler Intelligence Scale for Children – Revised (WISC-R) each year from 6 to 10 years of age. Of 321 infants originally enrolled in the study, 175 were tested after 5 years of age, and only 150 had complete data and were included in the analyses. The study authors used linear mixed models with random intercept and slope were used to analyze the pattern of BLL effects on Full Scale IQ (FSIQ) from 6 to 10 years of age. Fixed effects of the mixed models were maternal IQ, child's sex, SES, birth weight, geometric mean BLL from age 1 to 5, BLL at each age FSIQ measures were taken, geometric mean BLLs during the second and third trimesters of pregnancy, and a variable indicating the first FSIQ measurement of the child. Results of the mixed models showed that mothers with higher BLL during the third trimester had children with significantly lower FSIQ scores (β = -3.90, 95% CI: -6.45, -1.36), VIQ scores (β = -3.15, *p*=0.007), and PIQ scores (β = -4.37, *p*=0.004) at 6 to 10 years of age. Similar analyses including prenatal BLL measurements at 12, 20, 28, and 36 weeks of pregnancy, rather than the averages prenatal BLL, demonstrated that only BLL at week 28 of pregnancy was significantly associated with FSIQ scores (β = -4.13, 95% CI: -6.45, -1.81). Additionally, although mixed models were fit with the 6-month HOME score, this covariate score did not change the significance or magnitude of the effect sizes and therefore, it was omitted from the full model. The authors concluded that increased maternal blood lead concentration during the third trimester of pregnancy, especially around week 28, was associated with decreased intellectual development in children.

Bellinger et al. (1992) examined the association between childhood lead exposure and intellectual functioning at 10 years of age. A total of 249 infants born at Brigham and Women's Hospital in Boston, MA between August 1971 and April 1981 were enrolled in the prospective cohort study, however, only 148 children underwent cognitive testing at 10 years of age. The outcomes of interest were Full-Scale IQ scores on the Wechsler Intelligence Scale for Children-Revised (WISC-R) and Battery Composite scores on the Kaufman Test of Educational Achievement – Brief Form (K-TEA). Children's blood lead levels (BLLs) were measured at birth (cord BLL), 6, 12, 18, 24, and 57 months old, and at 10 years of age. Multiple regression models were used to estimate the association between child's lead exposure and cognitive performance. Models for all seven BLL measurement time points included HOME57 score (total HOME score measured at 57 months of age). Additionally, models included some combinations of the following variables: maternal age, family stress, child stress, race, birthweight, maternal IQ, number of daycare situations through 57 months, SES, marital status, sex, birth order, number of residence changes prior to 57 months, family balance, and parent's sense of competence. The mean BLLs for children at 6, 12, 18, 24, and 57 months, and 10 years of age were 6.7, 7.7, 7.8, 6.5, 6.3, and 2.9 µg/dL, respectively. After adjusting for confounders, there was a significant association between 24-month BLL and Full-Scale IQ (WISC-R) and Battery Composite (K-TEA) scores. Over the approximate range of BLL from 0 to 25 µg/dL, Full-Scale IQ was associated with a 5.8 point decline per 10 µg/dL increase in BLL at 24 months (95% CI: 1.7, 9.9; p=0.007) and Battery Composite score was associated with a 8.9 point decline per 10 µg/dL increase in BLL at 24 months (95% CI: 4.2, 13.6; p=0.0003). The authors concluded that elevated BLLs at 24 months of age were associated with intellectual deficits at age 10.

Bellinger and Needleman (2003) re-analyzed data from their prospective cohort study in Boston (Bellinger et al. 1992) to examine the subset of 48 children whose blood lead level (BLL) never exceeded

10 µg/dL at birth or at 6, 12, 18, 24, 57, or 120 months of age. Characteristics of this subgroup were not described. Scores on the Wechsler Intelligence Scale for Children-Revised and Kaufman Test of Educational Achievement were the primary outcomes assessed at 120 months of age. Multiple regression models adjusted for each age at which BLL measure included some subset of the following covariates: maternal age, family stress, child stress, race, birthweight, maternal IQ, number of daycare situations through 57 months, SES, marital status, sex, birth order, number of residence changes prior to 57 months, family balance, and parent's sense of competence (Bellinger et al. 1992). The authors reported a significant inverse association between IQ at 120 months and BLL at 24 months of age after adjusting for unspecified covariates (p = 0.03). Further, nonparametric smoothing analyses indicated that this inverse association remained at blood lead levels below 5 µg/dL. The blood lead coefficient was -1.56, greater than previous analyses including with lead levels above 10 µg/dL (-0.58). The authors concluded that children's IQ scores are reduced without a threshold level.

Evens et al. (2015) conducted a retrospective cohort study to examine the association between lead concentration of Chicago Public School (CPS) children and their performance on 3rd grade Illinois Standard Achievement Test (ISAT) reading and math scores. The authors examined 57,350 children who were born in Chicago between 1994 and 1998, lived in Chicago during early childhood, had blood lead test results reported to the Chicago Department of Public Health (CDPH) between 1996 and 2006, and were enrolled in 3rd grade at a Chicago public school between 2003 and 2006. The dataset was created by linking the Chicago birth registry, Chicago Blood Lead Surveillance Program database, and CPS performance records (i.e., ISAT reading and math scores). Blood lead tests were used as a measure of childhood lead exposure and the primary outcome measures were 3rd grade ISAT math and reading scores, and additional outcomes included math score and reading score failure rates. The analyses were restricted to a sample of 47,168 children that had blood lead levels of < 10 µg/dL, and the mean blood lead level was 4.81 µg/dL (SD = 2.22). Multivariable linear regression models and multivariable log binomial models were used to assess the association between childhood blood lead levels ≤10 µg/dL and school performance, stratifying by race/ethnicity and adjusting for gender, mother's education, low-income (determined based on free or reduced-price lunch program enrollment), very low birthweight/early preterm, child's age at time of blood lead measurement, ISAT vs. Iowa exam type, and race. For children of all races/ethnicities, there was a significant decline in reading and math scores, respectively, per unit increase in blood lead concentration was -0.60 (SE = 0.03; p = <0.0001) and -0.50 (SE = 0.03; p<0.0001). The interaction terms between blood lead levels and race/ethnicity were statistically significant (p < 0.0001). Stratified by race/ethnicity, associations in reading and math declines, respectively, were steepest for non-Hispanic white children (reading: β = -0.75, p = <0.0001; math: β = -0.71, p = <0.0001), followed by Hispanic children (reading: β = -0.60, p = <0.0001; math: β = -0.52, p = <0.0001), and then non-Hispanic black children (reading: β = 0.57, p = <0.0001; math: β = -0.44, p = <0.0001). For children of all races/ethnicities, there was a statistically significantly increased risk of failing reading (RR = 1.32, 95% CI: 1.26-1.39) and math (RR = 1.32, 95% CI: 1.26-1.39) associated with each 5 µg/dL increase in blood lead level. The highest associations between an increased blood lead level of 5 µg/dL and the risk of failing reading and math were observed among non-Hispanic white children (reading RR = 1.93, 95% CI: 1.47-2.54; math RR = 1.71, 95% CI: 1.26-2.30), Hispanic children (reading RR = 1.47, 95% CI: 1.29-1.66; math RR = 1.51, 95% CI: 1.31-1.75), and non-Hispanic black children (reading RR = 1.28, 95% CI: 1.21-1.35; math RR = 1.28, 95% CI: 1.22-1.35). The authors concluded that their study demonstrated that small increases of blood lead levels, even those <5 µg/dL, were significantly associated with decrements in performance on standardized tests.

Blackowicz et al. (2016) conducted a retrospective cohort study of Hispanic children in Chicago, and linked data from the Chicago birth registry, Chicago Blood Lead Registry, and 3rd grade Illinois Standard Achievement Test (ISAT) scores in order to examine the association between blood lead levels (BLL) and standardized test performance by the following Hispanic subgroups: Mexican, Puerto Rican, and Other Hispanic. The sample included 12,319 Hispanic children that were born in Chicago between 1994 and 1998, lived in Chicago during early childhood, were enrolled in 3rd grade at a Chicago public school between 2003 and 2006, had a BLL test reported to the Chicago Department of Public Health between 1996 and 2006, and had a BLL level less than 10 µg/dL. The mean BLL for this sample of children was 4.16 µg/dL and the median BLL was 4.0 µg/dL. For children with more than one reported BLL (less than 10% of the sample), the most recent BLL was used. The primary outcome variable was 3rd grade ISAT score, however additional outcome variables included math score failures and reading score failures as defined by the ISAT score. Ordinary least squares regression was used to assess crude relationships between BLL and ISAT scores, and log binomial regression was used to assess crude relationships between BLL and failure rate for math and reading sections. Multivariable models adjusted for gender, mother's education level, low-income (defined by subsidized vs. unsubsidized school lunches), small for gestational age (SGA), preterm birth, child's age at time of BLL, exam type (ISAT vs. Iowa), and Hispanic subgroup. There was no significant interaction between Hispanic subgroup and BLL for any of the models, and therefore this interaction term was excluded from the final models. After adjustment for confounders, there was a significant inverse association between BLL and ISAT scores, specifically, each 1 µg/dL increase in BLL was associated with lower average reading scores by 0.55 points ($p<0.001$) and lower average math scores by 0.48 points ($p<0.001$). In addition, reading and math failure rates were significantly higher in children with higher BLL, specifically reading failure rates were 43% higher for each 5 µg/dL increase in BLL (RR = 1.43; 95% CI: 1.25, 1.63) and math failure rates were 53% higher for each 5 µg/dL increase in BLL (RR = 1.53; 95% CI: 1.32, 1.78). The authors concluded that among Hispanic children in Chicago public schools with BLLs below 10 µg/dL, higher BLLs are associated with an increased risk for poor performance on standardized reading and math tests in 3rd grade.

Min et al. (2009) conducted a prospective cohort study to evaluate the association between lead exposure at 4 years of age and children's IQ and academic achievement at 4, 9, and 11 years of age, among urban, poor, and prenatally poly-drug exposed children. Inner-city children from Cleveland, OH, were recruited at 4 years of age, and included women who used alcohol, marijuana, or tobacco during pregnancy. Of the 278 children in the sample, 86% were African American, 98% came from a low socioeconomic background, 88% were prenatally exposed to at least one substance, 51% were prenatally exposed to cocaine, and 77% were prenatally exposed to alcohol. An abbreviated Wechsler Preschool and Primary Scales of Intelligence-Revised was administered to children at 4 years of age produced scores for Full Scale, Verbal, and Performance IQ. The Wechsler Intelligence Scales for Children – Fourth Edition and the Woodcock Johnson-III Tests of Achievement administered to children at both 9 and 11 years of age resulted in Full Scale IQ, verbal comprehension, perceptual reasoning, working memory, processing speed, math, and reading scores. Linear regression models for all outcomes adjusted for HOME score, as well as some subset of the following covariates that varied for each model: maternal or current caregiver's Peabody Picture Vocabulary Test score, child sex, race, parity, maternal marital status, infant head circumference, iron deficiency anaemia (IDA), prenatal cocaine, maternal years of education, maternal or current caregiver WAIS-BD score, prenatal alcohol, current caregiver WAIS-PC score, prenatal marijuana, and current caregiver alcohol use. Mean blood lead level (BLL) measured at 4 years of age was 7 µg/dL, and ranged from 1.3 to 23.8 µg/dL. An estimated 4.1 to 5.4 Full Scale IQ decrement was significantly associated with each 10 µg/dL increase in blood lead level at ages 4, 9, and 11 years ($p<0.05$). Significant associations were additionally observed between BLL at age 4 and Performance IQ at age 4 ($\beta = -0.74$, $p<0.001$), perceptual reasoning at age 9 (β

= -0.45, p<0.05) and age 11 (β = -0.61, p<0.01), reading scores at age 9 (β = -0.58, p<0.05) and age 11 (β = -0.60, p<0.01), and verbal comprehension (β = -0.51, p<0.01) and math scores at age 11 (β = -0.45, p<0.05). Further analyses of a subgroup of children with BLLs <10 µg/dL showed children with BLLs from 5 to <10 µg/dL had significantly worse Performance IQ at age 4 (p=0.01), perceptual reasoning at age 9 (p=0.01), and reading scores at age 9 (p=0.003) and 11 (p=0.04) compared to children with BLLs < 5 µg/dL. The authors concluded that early lead exposure effects cognitive outcomes and school achievement through late childhood.

Miranda et al. (2007) conducted a retrospective cohort study that aimed to investigate whether blood lead levels in early childhood were related to educational achievement in early elementary school, based on performance on End-of-Grade testing, in the Piedmont region of North Carolina. Blood lead surveillance data was available from the North Carolina Childhood Lead Poisoning Prevention Program registry and educational testing data was available from the North Carolina Education Research Data Center (NCERDC). Children who had blood lead level (BLL) data and were screened for lead from 0 to 5 years of age during 1995 to 1998 and were subsequently matched to their 4th grade End-of-Grade test data. Linear regression models controlled for sex, race, participation in the free or reduced price lunch program, parental education, daily computer use, whether school is a charter school, age at which blood lead screen occurred, and dummy variables for each of the school systems. Among children with blood lead data, there were 8,603 subjects in the reading data set and 8,627 subjects in the math data set. Mean BLL was 4.52 and 4.53 µg/dL for the reading and math data sets, respectively. Three different models used different constructions of blood lead exposure: BLL as a continuous variable, with dummy variables for BLL between 5-9 µg/dL and ≥10 µg/dL, and with dummy variables for BLLs of 2, 3, 4, …, 9, and ≥10 µg/dL. The coefficient for 4th grade reading score data was associated with a BLL of 2 µg/dL and marginally significant at p = 0.05, the rest of the coefficients between reading and math score data with BLLs were significant at p<0.05. The authors concluded that the relationship between blood lead levels and cognitive outcomes were robust, and that early childhood lead exposure reduced test scores even at low levels of lead exposure.

Miranda et al. (2009) carried out a retrospective cohort study in an interest to confirm earlier results in a larger sample of the population (Miranda et al. 2007), determine whether there were any differences in the impact of lead across the distribution of the End-of-Grade (EOG) test scores, and to clarify the impact of cumulative childhood social and environmental stress on educational outcomes. Blood lead surveillance data was available from the North Carolina Childhood Lead Poisoning Prevention Program registry and educational testing data was available from the North Carolina Education Research Data Center (NCERDC). Analyses were expanded to include children from all 100 counties of North Carolina that had blood lead level (BLL) data from a lead screening at 9 to 36 months of age between the years 1995 and 1999, and 4th grade EOG reading test score data from age-corresponding years. Multivariate and quantile regression analyses adjusted for sex, race, household income, parental education, and individual school system variability. The total sample included 57,678 children, of whom 43.1% were black and 56.9% were white, and 52.9% were enrolled in the free or reduced lunch program. BLLs of the sample ranged from 1 to 16 µg/dL and had a mean of 4.8 µg/dL and median of 4.0 µg/dL. BLL was included in the models as indicator variables for BLLs of 2, 3, 4, … 9, and ≥10 µg/dL, with a BLL of 1 µg/dL as the reference level. Each indicator variable for BLL was significantly (p<0.05) associated with decrements on EOG reading scores, and the estimated decreases in EOG reading scores became larger as BLL increased (β = -0.3 for a BLL of 2 µg/dL, and β = -1.75 µg/dL for BLL ≥10 µg/dL) indicating a dose-response relationship. Results of the quantile regression analyses showed that higher BLLs spread the lower tail of the distribution of EOG reading test scores more than the middle or upper tail of the distribution, and that the differential effect on the distribution of scores is significant (p = 0.04). Across

the distribution of reading test scores, parental education accounted for 58-65% of the total decrements, enrolment in the lunch program accounted for 25-28%, and BLL accounted for 7-16%. The authors concluded that early childhood lead exposure, as well as low parental education and poverty, were associated with decrements End-of-Grade test scores.

McDermott et al. (2011) assessed whether soil concentrations of metals in proximity to maternal residences during pregnancy were associated with probability of intellectual disability (ID) among children in a retrospective cohort study. Mothers were insured by South Carolina Medicaid between 1996 and 2002, and lived in one of six residential strips (four small towns, two small cities) during their sixth month of pregnancy. Medicaid reimbursement files for mothers and children, birth certificates, and children's school and agency records through May 2008 were merged, resulting in 8 to 12 years of follow-up time. Soil samples were taken in the six strips where pregnant women resided, and geocoding of maternal residences and Bayesian Kriging models estimated residential soil metal concentrations at residential addresses. Lead concentrations in the soil ranged from 0.9 to 1800 mg/kg dw, with a mean of 35.4 mg/kg dw and median of 19.0 mg/kg dw. The outcome measure was a confirmed diagnosis of ID with an unknown cause, which was determined using ICD9 codes obtained from Medicaid records and a record of eligibility determination from an agency providing services to children with ID. The authors initially looked at Spearman rank correlations between the eight metal concentrations and found the highest correlation (0.603) was for Arsenic (As) and Lead (Pb). Generalized additive models (GAM) were used to estimate the associations between the concentration of As and each metal with the risk of ID. They identified effect modification for the relationship between infants that were small for gestational age (SGA) vs. normal weight, and stratified the analysis based on SGA vs. normal weight. The models included an interaction term for As and Pb, started with a full model including all covariates, and then used backwards selection to form the final models for SGA and normal weight children. The covariates included in the final model were gestational age, child's sex, parity, child's age at last follow-up, and the interaction between As and Pb. It was reported that there was a significant association between Pb and ID (OR = 1.002; 95% CI: 1.000, 1.004) for children that were normal weight. In addition, the interaction between As and Pb was significant for normal weight children (p = 0.019). The authors concluded that there was a significant association between Kriged concentrations of soil Pb proximal to maternal residences and ID among normal weight children.

Longitudinal Studies of Lead and Cognition

Fergusson et al. (1988) conducted a longitudinal cohort study of New Zealand children in order to examine the relationship between lead body burden and cognitive ability. The cohort included children born in the Christchurch, New Zealand, urban region in mid-1977 that were studied at birth, 4 months of age, and annually until 9 years of age. Dentine lead levels were measured by shed deciduous teeth that children provided from the ages of 6 to 9 years of age and cognitive ability was assessed through concurrent administration of the Wechsler Test of Child Intelligence (WISC-R) and Burt Reading Test at 8 and 9 years of age as well as through annual teacher ratings of school performance. The original cohort consisted of 1265 children, however, due to missing data samples sizes for the analyses included: 1) 724 children for whom complete data on dentine lead values, IQ and word recognition tests, and all covariates were available at 8 years; 2) 886 children for whom data on dentine lead values, teacher ratings, and all covariates were available at 8 years; 2) 664 children for whom complete data on dentine lead values, IQ and word recognition tests, and all covariates were available at 9 years; or 4) 853 children for whom data on dentine lead values, teacher ratings, and all covariates were available at 9 years. It was noted that there was a tendency for children who were more disadvantaged at birth to be less likely to be included in the analysis samples. No further details were reported regarding cohort characteristics. The mean

dentine lead level for the sample was just over 6 µg/g (exact mean value was not reported). Multiple regression analyses solved with LISREL modelling methods were performed for each outcome variable that accounted for test reliability, sample selection factors, and covariates that made significant contributions to the outcome, which included maternal and paternal education, birth order, ethnicity, birthweight, gender, standard of living, maternal emotional responsiveness, maternal avoidance of punishment, number of schools attended, attendance at preschool education, and number of old weatherboard homes lived in. Results of the multiple regression analyses that accounted for test reliability, controlled for selection bias, and adjusted for covariates showed that there were no significant associations between dentine lead levels and Verbal IQ, Performance IQ, or Total IQ at 8 or 9 years of age. However, significant associations were observed between dentine lead levels and Burt reading test score at age 8 ($\beta = -0.08$, $p<0.05$) and age 9 ($\beta = -0.10$, $p<0.01$), as well as teacher performance ratings of reading at age 8 ($\beta = -0.13$, $p<0.001$) and 9 ($\beta = -0.08$, $p<0.05$), written expression at age 8 ($\beta = -0.13$, $p<0.001$) and 9 ($\beta = -0.08$, $p<0.05$), spelling at age 8 ($\beta = -0.14$, $p<0.001$) and 9 ($\beta = -0.09$, $p<0.05$), math at age 8 ($\beta = -0.08$, $p<0.05$) and 9 ($\beta = -0.10$, $p<0.05$), and handwriting at age 8 ($\beta = -0.12$, $p<0.01$) and 9 ($\beta = -0.14$, $p<0.01$). The authors concluded that the results of the study support the hypothesis that low level lead exposure may have negative effects on levels of achievement in children.

Fergusson et al. (1993) conducted a prospective cohort study of New Zealand children in order to assess the relationship between dentine lead levels at 6 to 8 years of age with cognitive and behavioral outcomes at 12 to 13 years of age. A total of 1265 children were initially enrolled in the cohort, however the sample sizes for the analyses ranged from 690 to 891 due to missing data. No additional cohort characteristics were reported. In the current study, dentine lead levels were measured when children were 6 to 8 years old and the mean dentine lead level for the sample was 6.2 µg/g (SD=3.7). Between the ages of 12 to 13, behavioral and cognitive outcomes were measured using the New Zealand revision of the Burt Reading Test administered, the Progressive Achievement Test (PAT) of reading comprehension, the Test of Scholastic Abilities (TOSCA), teacher ratings of school performance in the areas of reading, writing, and math, and parental and teacher responses to Rutter and Conners questionnaires measuring inattention and restlessness. Statistical analyses included multiple linear regression models to adjust for potential confounders while estimating the association between dentine lead levels and outcomes scores, and latent variable models based on LISREL methods to address potential measurement errors. Confounders that were adjusted for included gender, ethnicity, family size, maternal education, paternal education, SES, maternal emotional responsiveness (based on HOME inventory), avoidance of punishment (based on HOME inventory), number of schools attended, and residence in old weatherboard housing. For the analyses, dentine lead level was categorized into five groupings: 0-2, 3-5, 6-8, 9-11, and ≥12 µg/g. After adjustment for confounders, there were significant associations between dentine lead level and Burt reading score at age 12 ($\beta = -0.07$, $p<0.05$); TOSCA score at age 13 ($\beta = -0.07$, $p<0.05$); teacher ratings of reading at age 12 ($\beta = -0.10$, $p<0.005$) and 13 ($\beta = -0.07$, $p<0.05$), written expression at age 12 ($\beta = -0.11$, $p<0.001$) and 13 ($\beta = -0.08$, $p<0.01$), and math at age 12 ($\beta = -0.06$, $p<0.05$) and 13 ($\beta = -0.16$, $p<0.001$); and inattentive/restless behaviors at age 12 ($\beta = 0.11$, $p<0.001$) and 13 ($\beta = 0.06$, $p<0.05$). Results of the multivariate regression model using LISREL modeling methods rejected the null hypothesis that lead level was unrelated to all the test outcomes ($p<0.001$). The latent variable models correcting for measurement errors produced consistent results. The authors concluded that the results were consistent with the hypothesis that early mildly elevated lead levels may have long-term consequences for children's behavioral and cognitive development, however they additionally noted that observed associations between lead levels and the outcomes studied were small.

Fergusson et al. (1997) conducted a prospective cohort study of New Zealand children in order to examine the association between dentine lead levels at 6 to 8 years of age with educational success at

18 years of age. Out of the initial cohort of 1265 children, a total of 881 children enrolled in the Christchurch Health and Development Study that were evaluated at birth, 4 months of age, annually until 16 years of age, and again at 18 years of age were included in the current study. Dentine lead levels were measured from a shed deciduous tooth that children submitted between 6 and 8 years of age and the mean dentine lead level for the sample was 6.2 µg/g. Outcome measures used to assess cognition and educational success at age 18 included the New Zealand revision of the Burt Word Reading Test score, the proportion of children below a 12-year-old reading level (as measured by performance on Burt Reading Test), the proportion of children that failed to complete 3 years of secondary school, the proportion of children that left school without formal qualifications, and the number of School Certificate (series of national examinations administered to New Zealand students) passes. Multiple linear regression and multiple logistic regression models were used to estimate the association between dentine lead levels in childhood and educational performance at 18 years of age. Models for each outcome measure adjusted for covariates that were significantly associated with either dentine lead levels or the outcome measure, and included some subset of the following variables: gender, maternal age, maternal education, socioeconomic status, standard of living, breastfeeding duration, birth order, parental conflict, maternal punitiveness, school class level, and years lived near busy roads. For the analyses, dentine lead level was categorized into 5 levels: 0-2, 3-5, 6-8, 9-11, and ≥12 µg/g. After adjustment for potential confounders, there were significant associations between dentine lead levels at 6 to 8 years of age with educational outcomes at 18 years of age. Specifically, children with higher lead levels had lower reading test scores ($p<0.002$), more often had a reading level below that of a 12-year-old ($p<0.001$), more often failed to complete 3 years of secondary school ($p<0.02$), more often left school without educational qualifications ($p<0.05$), and had lower numbers of School Certificate passes ($p<0.05$). Population attributable risks (PARs) were calculated for dichotomous outcomes, which ranged from 4% to 12% (the specific PAR for each outcome was not reported). Secondary analyses additionally adjusted for measures of educational achievement at 12 to 13 years of age in order to explore whether school achievement in childhood/early adolescence mediated the association between dentine lead levels and later educational outcomes at age 18. Results of the regression models that additionally adjusted for educational performance at age 12-13 eliminated the significance of the associations between dentine lead levels and all educational outcomes assessed at age 18. The authors concluded that children with early elevated lead levels were at risk for educational difficulties extending to later adolescence. In addition, they conclude that their results suggest that early dentine lead levels were linked with educational outcomes in later adolescence through a causal chain model in which early dentine levels were associated with reduced educational performance in middle childhood and adolescence, which then leads to decreases in educational achievement in young adulthood.

Huang et al. (2012) conducted a prospective cohort study to understand the relationship between prenatal and postnatal exposure to low levels of lead and neurodevelopmental effects, including cognitive function and IQ, in young children. A total of 430 pregnant women were initially recruited from a medical center in Taichung, Taiwan, from 2001 to 2002, blood samples of 323 paired women and children were obtained at delivery, and their children were followed from 2003 to 2009 (119 were followed up at 2-3 years of age, 76 at 5-6 years of age, and 66 at 8-9 years of age). The authors analyzed maternal blood samples taken during the 3rd trimester, cord blood from newborns at delivery, and samples taken from children at 2-3, 5-6, and 8-9 years of age to determine blood lead levels. In order to evaluate intellectual development, the authors used Bayley Scales of Infant Development – II (BSID-II) to evaluate children at the 2-3 years of age follow-up, the Chinese version of Wechsler Preschool and Primary Scale of Intelligence – Revised (WPPSI-R) to evaluate children at the 5-6 years of age follow-up, and the Chinese version of Wechsler Intelligence Scale for Children-version III (WISC-III) to evaluate

children at the 8-9 years of age follow-up. Outcomes of interested included the mental development index (MDI) and psychomotor development index (PDI) measured at age 2-3 with the BSID-II, $FSIQ_5$ (full scale IQ) measured at age 5-6 with the WPPSI-R, and $FSIQ_8$ measured at age 8-9 with the WISC-III. The geometric mean blood lead levels for 3rd trimester maternal blood samples, cord blood samples, and at 2-3 years, 5-6 years, and at 8-9 years, respectively, were 2.21, 1.30, 2.48, 2.49, 1.97 µg/dL. Multivariate multiple regression models were used to assess the effects of blood lead level on IQ score at each age. These models adjusted for gender, maternal education, maternal age, maternal 3rd trimester blood lead level, and maternal alcohol consumption during pregnancy. A linear mixed model was used to examine the association between children's blood lead levels at 2-3 and 5-6 years of age with IQ scores at 5-6 and 8-9 years of age. Three linear mixed models were fit, the first model adjusted for gender, HOME score, and birth weight; the second model adjusted for the same covariates as the first model as well as maternal blood lead level, maternal education, and maternal age; the third model adjusted for the same covariates as the second model in addition to pesticide use, pre-pregnancy smoking, and post-pregnancy alcohol drinking. Results of the multivariate multiple regression models showed that there was a significant inverse association between children's log blood lead levels at 5-6 years of age and their IQ score at 8-9 years of age (β = -0.313, p = 0.012). In addition, marginally significant inverse associations were observed between children's log blood lead level at age 2-3 and their IQ at age 5-6 years (β = -0.21, p = 0.082) and their IQ at age 8-9 years (β = -0.22, p = 0.071), and between children's log blood lead level at age 8-9 years and their IQ at age 8-9 years (β = -0.219, p = 0.074). Results of all three of the linear mixed models showed that postnatal log blood lead levels at ages 2-3 years and 5-6 years were significantly associated with decreases in IQ scores at ages 5-6 years and 8-9 years (β = -6.56, p = 0.014; β = -5.69, p = 0.030; β = -5.97, p = 0.025; for models 1 to 3, respectively). No significant association was observed between prenatal and cord blood lead levels and cognitive outcomes for children ages 2 to 8 years old. The authors concluded that low-level postnatal lead exposure in children at 2-5 years of age may have lagged effects on children's IQ at 5 to 9 years of age.

Mazumdar et al. (2011) conducted a prospective cohort study in order to assess the association between childhood environmental lead exposure and intellectual functioning in adulthood as well as identify which time period blood-lead concentrations were most predictive of late outcome. Participants in the current study were recruited at 28 to 30 years of age from a cohort that was initially enrolled as infants from Brigham and Women's Hospital in Boston, MA, between August 1979 and April 1981. In the initial follow-up, blood lead levels (BLLs) were measured in cord blood, at 6, 12, 18, 24, and 57 months, and at 10 years of age, and the mean (SD) BLL at each interval was 6.5 (5.3), 8.0 (5.3), 10.0 (6.7), 7.7 (4.0), 6.7 (3.6), and 3.0 (2.7) µg/dL, respectively. Of the 148 children that were followed through 10 years of age, 143 were eligible as adults, and a total of 43 submitted a blood sample and participated in neuropsychological testing. The primary outcome of interest used to measure intellectual function in adulthood was Full-Scale IQ tested via the Wechsler Abbreviated Scale of Intelligence (WASI), although Verbal IQ and Performance IQ were also considered. The sample had a mean age of 29 years, were mostly college graduates (81.4%), white (93%), and had mothers that graduated college (60%). Separate, unadjusted linear regression models were fit for Full-Scale IQ with each lead exposure variable (i.e., blood lead levels taken at each age, and as an average, maximum, average early childhood, and average late childhood). In unadjusted models, average late childhood blood lead concentration (means of 4 and 10 year blood lead concentrations) had the strongest relationship with Full-Scale IQ (β = -1.89; 95% CI: -3.3, -0.47; p = 0.01; R2 = 0.154), although Full-Scale IQ was significantly related to blood lead concentrations at 6 months (β = -0.90; 95% CI: -1.73, -0.77; p = 0.03; R2 = 0.117), 4 years (β = -1.26; 95% CI: -2.43, -0.09; p = 0.04; R2 = 0.110), and 10 years (β = -1.95; 95% CI: -3.56, -0.33; p = 0.02; R2 = 0.145). Due to small sample size, covariates considered were only included in the models of lead and IQ one at a

time and included the following: HOME score, gender, birthweight, birth order, gestational age, mother's marital status at time of delivery, mother's education, maternal IQ, race, maternal smoking during pregnancy, maternal alcohol use during pregnancy, history of concussion, subject's smoking. Only the inclusion of maternal IQ in the model reduced the average late childhood blood lead index (β = -1.11; 95% CI: -2.29, 0.06) and changed the significance of the regression coefficient. The authors concluded that the results of their study suggested that lead exposure during childhood impacts intellectual functioning in young adulthood and that school-age lead exposure may be a period of increased susceptibility, although they noted that the potentially confounding effects of maternal IQ could not be excluded.

Taylor et al. (2017) is the report of a prospective birth cohort study from the United Kingdom (UK) that examined the relationship between prenatal lead exposure and child IQ at ages 4 and 8. The study population represents a sample of women and children from "ALSPAC" a population-based study investigating environmental and genetic influences on health, behavior, and development of children. The ALSPAC study recruited 14,541 pregnant women with an expected delivery date between April 1, 1991 and December 31, 1992. There were 14,062 live births associated with these women. Approximately 4285 study participants (women) had blood samples drawn at 11 weeks (between 9 and 13 weeks). A randomly selected sample of these women were invited to participate in this study of prenatal blood levels and IQ (n = 582). In the final analysis, there were 348 children evaluated for IQ at age 4 and 1826 children evaluated for IQ at 8 years. The IQ measures were verbal IQ, performance IQ, and full-scale (total) IQ. Information collected on confounders included family adversity index, housing tenure, household crowding, smoking in the first trimester, alcohol consumption in the first trimester, maternal age at index birth, parity, maternal education, length of time spent in health district, child sex, child age at testing, weighted life events score, and hemoglobin level. Results revealed a mean prenatal blood lead level of 3.67 ± 1.46 µg/dL. Fourteen percent (14%) of women had blood lead levels greater than or equal to 5 µg/dL. The mean child lead level was 4.22 ± 3.12 µg/dL with 26.6% having levels ≥ 5 µg/dL. The authors found "no evidence for any differences in 4-year IQ scores by prenatal lead category" (Taylor et al. 2017, p. 165).

Wasserman et al. (1997) conducted a prospective cohort study in order to examine the association between lead exposure and intelligence of school-aged children. The study recruited women who were pregnant from 1984 to 1985 and living in one of two towns in Kosovo, Yugoslavia: Mitrovica, the site of a lead smelter, or Pristina, located 25 miles away. Blood lead levels (BLLs) were measured mid-pregnancy and then at 6-month intervals from birth until approximately 7 years of age. In a total of 309 children, intelligence was assessed at 5 years of age using a translated version of the Wechsler Preschool and Primary Scale of Intelligence-Revised (WPPSI-R) and at 7 years of age with a translated version of the Wechsler Intelligence Scale for Children-Version III (WISC-III). Outcomes of interest derived from these tests included Performance IQ, Verbal IQ, Full Scale IQ, Perceptual Organization, Verbal Comprehension, and Freedom from Distractibility scores. BLL was analyzed as a cumulative lifetime lead exposure variable that was calculated as the area under the curve of log-transformed BLL measurements taken from birth until the time of intelligence testing, which resulted in a variable of cumulative exposure up until age 5 (AUC5) and a variable for cumulative exposure through age 7 (AUC7). Mean mid-pregnancy BLL, cord BLL, and cumulative BLL, respectively, was 13.3, 14.5, and 1.21 ug/dL. Ordinary least squares regression analyses were used to estimate the association between cumulative lifetime lead exposure with each of the outcomes of interest, adjusting for HOME score, maternal age, gender, sibship size at time of IQ test, birthweight, language spoken, years of maternal education, and maternal Raven's test score. After adjustment for covariates, there were significant associations between lifetime cumulative lead exposure (AUC7) and Full-Scale, Verbal, and Performance IQ scores at age 7 (all p's <0.001), as well

as Freedom from Distractibility ($p<0.01$), Perceptual Organization ($p<0.001$), and Verbal Comprehension ($p<0.05$) scores at age 7. A change in lifetime BLL (AUC7) from 10 to 30 µg/dL was associated with an estimated decrease in Full Scale, Verbal, and Performance IQ scores at age 7, respectively, of 4.3 points (95% CI: 3.4, 5.1), 3.4 points (95% CI: 1.7, 5.0), and 4.5 points (95% CI: 2.7, 6.3). The authors concluded that there were small, but statistically significant, associations between lead exposure and intelligence in school-aged children. In addition, they noted that cognitive functions related to perceptual organization and resistance to distractions seem more negatively affected by lead exposure than functions like language comprehension.

Wasserman et al. (2000) conducted a follow-up of Wasserman et al. (1997) to longitudinally assess the associations between the timing of lead exposure with IQ during preschool and early childhood years in Yugoslavian children. Women pregnant between 1985 and 1986 and living in one of two towns in Kosovo, Yugoslavia: Mitrovica (site of lead smelter) and Pristina (25 miles from smelter), were recruited for the study. Blood lead levels (BLLs) were measured at mid-pregnancy, at the time of delivery, and in children at subsequent 6-month intervals until age 7. Children's intelligence was assessed at age 3 and age 4 using the McCarthy Scales of Children's Abilities (MSCA), at age 5 with the Wechsler Preschool and Primary Scale of Intelligence-Revised (WPPSI-R), and at age 7 with the Wechsler Intelligence Scale for Children-Version III (WISC-III). A total of 390 children that received an intelligence test at age 3, 4, 5, or 7 and had complete covariate data were included in the present analyses. BLLs were $\log_{10}$ transformed and the mean $\log_{10}$ prenatal BLL, 0-2 year BLL, 2-3 year BLL, 2-4 year BLL, 2-5 year BLL, and 2-7 year BLL, respectively, were 1.01, 1.12, 1.28, 1.29, 1.28, and 1.24 µg/dL (equivalent to BLL ranging from 10.2 to 19.5 µg/dL). A prenatal lead exposure variable was calculated as the mean of the mid-pregnancy and delivery BLLs. A categorical variable characterizing patterns of change in postnatal lead exposure was constructed. To construct the postnatal change variable, early postnatal BLL was calculated as the mean of BLLs from birth to age 2 and late postnatal BLL was calculated as the mean of BLLs from 2.5 years old until the time of each IQ test (3, 4, 5, or 7 years old). Four categories were created that represented an increase (I) or no increase (NI) in early and late postnatal BLL compared to prenatal BLL. An increase was defined as a 50% or greater increase in postnatal BLL relative to prenatal BLL. The categories constructed were NI/NI, NI/I, I/NI, and I/I representing either no increase or increase for the early/late postnatal periods, respectively. The I/NI group was not included in the analyses as only 2 children fell in this category. Once the lead exposure variables were created, repeated measures linear regression was used to estimate the association between prenatal BLL and postnatal changes in BLL with intelligence, and model parameters were estimated with generalized estimating equations (GEE). For the postnatal change variable, the NI/NI category was used as the reference level. The model included the following covariates: maternal age, HOME score, child's age at time of intelligence test, child's sex, birthweight, ethnicity, years of maternal education, maternal Raven's score, and number of previous live births. After adjustment for covariates, there was a significant decrease of 6.05 IQ points (SE: 1.35, $p<0.001$) for each log unit increase in prenatal BLL, representing a 1.07 IQ point decrease (95% CI: 0.60, 1.53) for each 50% increase in prenatal BLL. Adjusting for covariates and prenatal BLL, there was a significant decrease in IQ associated with postnatal BLL increases occurring in either the late and/or early postnatal periods ($p<0.05$). A 50% increase in both early and late postnatal BLL relative to prenatal BLL was associated with a 2.71 IQ point decrease (95% CI: -4.91, -0.52), and a 50% increase in only the late postnatal BLL relative to prenatal BLL was associated with a 1.78 IQ point decrease (95% CI: -3.51, -0.05). The authors concluded that increases in prenatal and postnatal BLLs were independently associated with small decreases in young children's intelligence.

Studies with Cognitive and Behavioral Outcomes

Moodie et al. (2013) conducted a secondary analysis of data from a 2001 cohort of first grade children (N = 602) that attended one of nine schools within 3.2km of the lead smelter in Torreon, Mexico. The goal of the analysis was to examine the combined influences of lead exposure and quality of home environment on cognitive and behavioral outcomes in first grade children. The average child blood level of the sample was 11.4 µg/dL (SD = 6.1). Structural equation modeling included latent variables on cognition, behavior, and home environment. Cognition and behavior were comprised of 4 neurocognitive and behavioral latent variables (reasoning, visual spatial, and teacher and parent rating for ADHD), which were measured by the WISC-RM digit span test, WISC-RM math test, math achievement test measuring performance on the 1st grade math curriculum, Conners teacher and parent report of ADHD, Peabody picture vocabulary test, figure match test, figure design test, and trail making correct test. In addition, the structural equation models accounted for home environment, which included maternal education and maternal support for school work and extracurricular activities. Results showed that child lead exposure had a significant negative association with home environment (B = -0.180, p<0.001), cognition (β = -0.147, p<0.001), and behavior (β = -0.113, p<0.05). Results of the mediation model were significant and showed that home environment mediates the effects of lead on behavior (β = 0.305) and cognition (β = 0.184, p-value not provided). The authors concluded that lead exposure was negatively related to cognition and behavior, but noted that the effect size was low. They additionally concluded that home environment had a greater mediation effect on the association between lead and behavior compared with the mediation effect of home environment on lead and cognition.

Naicker et al. (2012) evaluated the association between blood lead levels and socio-behavioural problems among young adolescents in the South African Birth to Twenty (Bt20) cohort. Children in the Bt20 cohort were born between April and June 1990, resided in Soweto-Johannesburg area for at least 6 months after birth, and have had 16 follow up visits from birth to 20 years of age. A total of 1,041 children had blood lead levels (BLLs) measured at 13 years of age and also completed the Youth Self Report (YSR) at 13 years of age. The sample included 487 males and 554 females and the socio-demographic characteristics were not significantly different between males and females. Geometric mean BLL at age 13 for the total sample was 5.2 µg/dL (95% CI: 5.05, 5.32), however the geometric mean BLL was significantly higher for boys than girls (6.0 vs. 4.5 µg/dL, p<0.001). Behavioral outcomes were assessed using the YSR and were separated into rule-breaking behaviors (15 questions) and aggressive behaviors (17 questions). Log-transformed BLLs were used in the analyses. Multiple regression analyses were performed for the total sample and were also stratified by sex. Analyses of the total sample showed that the rule-breaking behaviors "running away" (p = 0.02) and "stealing outside the home" (p = 0.03) were significantly associated with BLLs; these associations were not significant after stratification by gender. Analyses of the total sample showed that the aggressive behaviors "destroys own things" (p < 0.001) and "threatens others" (p = 0.05) were significantly associated with BLLs. After stratifying by gender, it was found that among boys there were significantly higher geometric mean BLLs for boys with positive responses vs. negative responses to the aggressive behaviors "destroys own things" (8.2 vs. 6.0 µg/dL, p = 0.01), "attacks people" (9.2 vs. 6.0 µg/dL, p = 0.03), and "loud" (7.0 vs. 6.0 µg/dL, p = 0.02). In addition, boys with positive responses to "argues a lot" had a significantly lower geometric mean BLL compared to boys with negative responses (5.6 vs. 6.2 µg/dL, p = 0.03). Among girls, there were no significant associations between BLLs and aggressive behaviors. Further multivariate analysis adjusted for socio-economic index, which was measured as a composite score of 10 items including asset ownership and access to electricity, flush toilets, and indoor water. After adjustment for socio-economic index, there was a significant association between "argues a lot" (B = -

0.13; 95% CI: -0.23, -0.02) and "attacking people" (B = 0.54; 95% CI: 0.09, 0.98) and blood lead levels. The authors concluded that higher blood lead levels were associated with anti-social and destructive behaviors among boys in their early teens; however, they commented that other factors that affect behavior, including home environment, were unknown and not be considered in the analysis.

Chandramouli et al. (2009) assessed the relationship between early lead exposure at levels less than 10 µg/dL and behavioral and educational outcomes in school-aged children, to determine whether there was evidence of a threshold of effect. Participants were recruited from the Avon Longitudinal Study of Parents and Children (ALSPAC) in the UK, which included children born between April 1991 and December 1992. A randomly selected sample of 10% of the population cohort were invited to participate in the study. Lead concentrations were measured for 582 participants; however, only 337 to 425 participants were used in the analyses due to missing outcome data. Children's blood lead levels (BLLs) were measured at 30 months of age, and behavioral and academic outcomes were measured at 7 and 8 years of age. 94% of children had blood lead levels below 10 µg/dL. Educational performance was assessed at 7 years of age using the national Standard Assessment Tests (SATs) provided to all children in mainstream UK education. Behavior was assessed at 7 years of age using the Strengths and Difficulties Questionnaire (SDQ) and at 8 years of age using the Development and Well-being Assessment (DAWBA), Anti-social Behaviour Interview, and the Test of Everyday Attention for Children (TEACh). Regression analyses adjusted for child's gender, maternal education, home ownership, maternal smoking, home facilities score at 6 months, paternal SES at time of pregnancy, Family Adversity Index, and parenting attitudes at 6 months. BLLs were analyzed as a log-transformed continuous covariate and also as a categorical variable with levels 0 to <2 (reference), 2 to <5, 5 to <10, and >10 µg/dL. When modeled as a continuous covariate, BLL was significantly associated with reading (OR = 0.64; 95% CI: 0.47, 0.86; p = 0.004), writing (OR = 0.61; 95% CI: 0.46, 0.82; p = 0.001), and spelling (OR = 0.55; 95% CI: 0.36, 0.83; p = 0.004) SATs scores and with antisocial behaviors (OR = 1.54; 95% CI: 1.01, 2.33; p = 0.04). When modeled as a categorical variable, children with BLLs from 5 to 10 µg/dL had significantly lower reading (OR = 0.51; 95% CI: 0.32, 0.82; p = 0.006) and writing (OR = 0.49; 95% CI: 0.31, 0.78; p = 0.003) SATs scores compared to children with BLLs from 0 to 2 µg/dL. Children with BLLs greater than 10 µg/dL had significantly greater antisocial behaviors (OR = 2.90; 95% CI: 1.05, 8.03; p = 0.04), teacher-reported hyperactivity (OR = 2.82; 95% CI: 1.08, 7.35; p = 0.034), and teacher-reported total difficulties (OR = 2.69; 95% CI: 1.06, 6.81; p = 0.037) compared to children with BLLs ranging from 0 to 2 µg/dL. It was further reported that doubling BLL exposure from 5 to 10 µg/dL was associated with a 0.2 (95% CI: -0.03, -0.8) point decline in SATs writing scores and a 0.3 (95% CI: -0.9, 0.63) point increase in teacher-reported hyperactivity scores. The authors concluded that early childhood lead exposure, even at low levels between 5 to 10 µg/dL, is harmful to subsequent educational attainment and behaviors.

Taylor et al. (2015) conducted a prospective birth cohort study of mothers and children in the UK in order to examine the associations between in utero lead and cadmium exposure and childhood lead exposure with children's balancing ability at 7 and 10 years of age. The study sample was obtained from the Avon Longitudinal Study of Parents and Children (ALSPAC), and it was reported that the social and demographic characteristic of the cohort were similar to those in the UK national census surveys. Prenatal blood lead levels were available for 4285 women and childhood blood lead levels were measured for 582 children at 30 months of age. A random selection of 10% of parents with babies born in the last 6 months of the ALSPAC study were invited to provide their child's blood samples. The mean child blood lead level was 4.22 µg/dL (SD = 3.12) and the mean prenatal blood lead level was 3.67 µg/dL (SD = 1.47). Outcome measures included the results of the heel-to-toe walking test of the Movement Assessment Battery for Children administered at 7 years; dynamic balance at 10 years as measured by a walking along a beam, heel-to-toe, eyes open test; static balance eyes closed at 10 years as measured by

the summation of scores from the head-to-toe balance on a beam, eyes closed test and the standing on one leg, eyes closed test. Additional outcomes include questionnaire items regarding balance abilities completed by the primary caregiver when the child was 18, 30, 42, and 81 months of age, as well as 10 years of age. Blood lead level was categorized into two groups: <5 µg/dL and ≥5 µg/dL. Unadjusted and adjusted logistic regression models were used to estimate the association between blood lead levels and balance outcome variables. Models for balance test outcomes at 7 and 10 years of age adjusted for sex, passive smoking, and calcium and iron intake. Models for balance questionnaire outcomes related to bicycle riding ability adjusted for traffic level on the home street, type of accommodation, lowest level of accommodation, and maternal education. Before and after adjustment for confounders, there were no significant associations between either maternal blood lead level or childhood blood lead level with balance test outcome at age 7, dynamic balance test outcome at age 10, or static balance score at age 10. Prenatal lead levels ≥5 µg/dL showed a significant association with not being able to ride a bike (p=0.007), however childhood lead levels ≥5 µg/dL were also associated with the ability to ride a bike very well without stabilizers (p=0.050). Neither of these associations were significant once the models were adjusted for confounders. The authors concluded that there was no evidence of an association between prenatal lead exposures or childhood lead exposure with balance ability in children at 7 and 10 years of age.

Leviton et al. (1993) conducted a prospective cohort study of children born in Boston, MA, between April 1979 and March 1980 in order to examine the associations between prenatal and postnatal lead exposures with learning problems and school functioning at 8 years of age. Of the 3,814 eligible babies that were screened for the study, a total of 2,657 babies were enrolled, and follow-up data was collected on 1,923 children that were described as a group of relatively privileged, mainly middle-class children. Umbilical cord blood lead levels (BLLs) measured at birth were used to represent prenatal lead exposure and dentine lead levels measured from a shed deciduous tooth at 6 years represented postnatal lead exposure. Mean umbilical cord BLL of the sample was 6.8 µg/dL and the median was 6.2 µg/dL. The mean dentin lead level of the sample was 3.3 µg/g and the median was 2.8 µg/g. During the end of the academic year when children were 8 years old, teachers completed the Boston Teacher Questionnaire (BTQ) to assess classroom behaviors and academic performance, which resulted in seven outcomes of interest, including behavior, hyperactivity, reading, arithmetic, directions, daydreaming, and tasks. Additional outcomes of interest included the teacher's responses to whether a child was functioning as well as their peers, whether a child received special services, and whether a child had repeated a grade. Logistic regression models were used to estimate the associations between both umbilical cord (prenatal) and dentine (postnatal) lead levels with each of the outcome measures related to school performance and behavior. Multivariate models were fit to explore the association between the outcomes with prenatal lead exposure while accounting for postnatal lead exposure, and vice versa. All models were carried out separately for boys and girls and all adjusted for the following dichotomous confounders: single-parent family, gestational age less than 37 months, mother not a college graduate, self-identification as black, only one child in the family, daycare during the first 3 years, and missing information about daycare. Results showed that there were significantly increased risk ratios related to elevated cord BLLs and the tasks cluster for girls (continuous RR = 2.3, 95% CI: 1.1-5.1; dichotomous RR = 2.1, 95% CI: 1.0-4.4) and the directions cluster for boys (continuous RR = 2.8, 95% CI: 1.3-6.0; dichotomous RR = 2.7, 95% CI: 1.4-5.2). Among girls, elevated measures of dentine lead significantly increased girls' risk of the reading cluster (continuous RR = 2.2, 95% CI: 1.1-4.2; dichotomous RR = 2.1, 95% CI: 1.1-4.2), tasks cluster (continuous RR = 3.7, 95% CI: 1.7-7.7; dichotomous RR = 2.3, 95% CI: 1.1-5.0), and dysfunction compared to peers (continuous RR = 2.5, 95% CI: 1.4-4.4; dichotomous RR: 1.7, 95% CI: 1.1-2.6). The authors concluded that, although most clusters do not appear to be associated

with lead exposure, their findings suggested that children may be at increased risk of selected clusters, even at lead levels that are only modestly higher than those typical of present-day community exposure.

Longitudinal Studies Involving Behavioral and Cognitive Outcomes

Winter and Sampson (2017) conducted a cohort study to evaluate the associations between childhood lead exposure and three domains of later adolescent health, including behavioral, mental, and physical health. Members of the original birth cohort from the Project on Human Development in Chicago Neighborhoods (PHDCN) were recruited from 1995 to 1997 and were followed until 17 years of age in the current study. Children's blood lead levels were tested four times by 36 months of age and additional testing was conducted until 6 years of age if the child moved or had previous high test results. Behavioral, mental, and physical outcomes measures were evaluated in 2012 to 2013, or approximately 15 years after blood lead levels were measured. The child's primary caregiver completed the Child Behavior Checklist to assess the behavioral mental health outcomes and physical health, as measured by the impulsivity scale, anxiety or depression scale, and body mass index (BMI), respectively. Statistical analyses included ordinary least squares regression models, adjusting for individual (gender and race/ethnicity), household (the primary caregiver's immigrant generational status, marital status, education level, and receipt of Temporary Assistance for Needy Families), and neighborhood characteristics (proportion of individuals by ethnicity and below the poverty line). Mean average blood lead level was 6.14 µg/dL among participants. The analyses included 208 subjects, which comprised 54% females and 46% males, a majority of whom were black or Hispanic. Minority and poor children were disproportionately exposed to lead compared to white and less poor children. Mean differences in blood lead levels between black children (7.48 µg/dL) and white children (4.86 µg/dL), and Hispanic children (4.59 µg/dL) and white children were statistically significant. Each 1.0 µg/dL of average blood lead was significantly associated with 0.06 (95% CI: 0.01, 0.12) standard deviations higher impulsivity scores, 0.09 (95% CI: 0.03, 016) standard deviations higher scores in anxiety or depression, and 0.37 (95% CI: 0.11, 0.64) points higher in BMI. The authors concluded that childhood lead exposure undermines adolescent well-being.

Ris et al. (2004) conducted a prospective cohort study to evaluate early exposure to lead and neuropsychological outcomes in mid-adolescence. Pregnant women were initially recruited in Cincinnati, OH, from 1979 to 1983, and those that were known to be addicted to drugs, alcoholic, diabetic, or with proven neurologic disorders, psychoses, or mental retardation were excluded. Furthermore, infants less than 35 weeks gestation at birth, less than 1,500 g at birth, and with defined genetic syndromes or other serious medical conditions at birth were ineligible for the study. Enrolled children underwent blood lead determinations, medical examinations, and developmental follow-up on a quarterly basis until 5 years of age, and again at 5.5, 6, 6.5, 10, and between 15 to 17 years of age. Of the 300 subjects that were initially enrolled, the present follow-up includes 195 children that were examined at 10 years and again between 15-17 years of age. Neuropsychological measures included memory, learning/IQ, attention, visuoconstruction and fine-motor tests. Multiple regression analyses were conducted that adjusted for maternal IQ, SES, average total HOME score, PbB x gender interaction, and PbB x SES interaction. Average blood lead levels in the current follow-up were not reported. However, Figure 1 shows average BLL for each follow-up visit by quartiles of exposed groups with group mean BLL ranging from 4 to 34 µg/dL. Prenatal blood lead was significantly associated with visuoconstruction (β = -0.157, p = 0.011) and attention (β = -0.156, p = 0.001); average childhood blood lead was significantly associated with attention (β = -0.113, p = 0.005); and 78 month blood lead was

significantly associated with fine motor (β = -0.046, p = 0.004). The authors concluded that the results indicated there was an association between early lead exposure and neuropsychological effects.

Section 5.3.2 Studies of Lead and Behavioral Disorders

Cohort Studies of Lead and Behavior in Infants, Toddlers, and School Age Children

Plusquellec et al. (2007) conducted a cohort study that examined the association between prenatal exposure to lead and several aspects of behavioral function during infancy among 169 Intuit infants from Quebec, Canada. Maternal blood samples at birth and cord blood samples were analyzed for blood lead levels (BLL), and the authors noted that cord BLL was used as the primary biomarker of exposure for their study. Average lead levels were 4.6 µg/dL and 5.9 µg/dL in cord and maternal samples, respectively. The BSID-II was administered at 11 months of age and the Behavioral Rating Scale (BRS) of BSID-II was analyzed as an outcome of interest, along with behavioral coding of video recordings taken during the administration of the mental development index (MDI) of the BSID-II. The authors stated that they did not analyze MDI or the psychomotor developmental index (PDI) because they had previously found there was no association between prenatal lead exposure and those endpoints in this cohort. Multiple regression was used with covariates varying by model; however, adjustment included variables such as language, IDESPQ depression scale, socioeconomic index, blood hemoglobin, maternal age, number of children in the home, and maternal alcohol consumption. The results indicated that after adjusting for confounders, cord BLL was significantly associated with frenetic movement (B= -0.16, p=0.03) as measured by the BRS, but not associated with any other BRS outcomes. It was reported that cord BLL was significantly associated with the off-task duration (B= 0.17, p=0.04) and off task latency (B= -0.20, p=0.02) measured via behavioral coding. The authors concluded that these results provide evidence that increasing the specificity and precision of behavioral assessments can improve our ability to detect low-to-moderate associations between BLL and infant behavior.

Liu et al. (2014) conducted a prospective cohort study to evaluate the association between blood lead concentrations and behavioral problems in Chinese preschool children. A total of 1,341 children aged 3 to 5 years were drawn from four preschools in Jintan city, Jiangsu province, China from fall 2004 to spring 2005. Blood lead concentrations were measured once for each child between the age of 3 to 5 years old and mean blood lead levels were 6.4 µg/dL (SD = 2.6). Behavioral problems were assessed at age 6 using the Child Behavior Checklist (CBCL/1.5-5) and the Caregiver-Teacher Report Form (C-TRF/1.5-5). Linear regression was performed to examine the relationship between blood lead concentration and behavior scores. Logistic regression was performed to examine the association between blood lead concentration and the odds of clinical-level behavioral problems. Both linear and logistic regression analyses adjusted for age at blood lead test, sex, preschool residence, father's education, mother's education, father's occupation, parents' marital status, single child status, and IQ. After adjustment, there were significant associations between blood lead concentrations and increased scores for teacher reported behavioral problems including emotional reactivity (β = 0.322; 95% CI: 0.058, 0.587), DSM-oriented anxiety (β = 0.253; 95% CI: 0.016, 0.500), and DSM-oriented pervasive developmental problems (β = 0.303; 95% CI: 0.046, 0.560). After adjustment, there were significantly increased odds between blood lead levels and various teacher-reported behavioral problems and syndromes, including internalizing problems (OR = 1.10, 95% CI: 1.03-1.18), emotionally reactive (OR = 1.10, 95% CI: 1.02-1.19), anxious/depressed (OR = 1.12, 95% CI: 1.03-1.23), DSM-oriented anxiety (OR = 1.10, 95% CI: 1.01-1.19), and DSM-oriented pervasive developmental problems (OR = 1.16, 95% CI: 1.07-1.25). The authors concluded that their study demonstrated that mean blood lead concentrations

of 6.4 µg/dL in children aged 5 to 6 years were significantly associated with increased behavioral problem scores at age 6 years.

Kim et al. (2016) conducted a prospective cohort study of 2473 children from 10 Korean cities in order to examine the association between blood lead levels and autistic behaviors among school children. Children were enrolled in the study in 2005-2006 when they were 7-8 years of age, examined again in 2007-2008 when they were 9-10 years of age, and followed up with a second time in 2009-2010 when they were 11-12 years of age. The sample of children in the study was 50.2% boys, 39.6% with a mother and 49.6% with a father that had more than a high school degree, and 34% had families with a monthly income less than $2000 USD. The geometric mean (95% CI) blood lead levels (BLLs) at ages 7-8, 9-10, and 11-12 were 1.64 (1.60, 1.68), 1.58 (1.55, 1.61), and 1.58 (1.55, 1.61) µg/dL, respectively. The Autism Spectrum Screening Questionnaire (ASSQ) and Social Responsiveness Scale (SRS) were completed by the parents when the child was 11-12 years old to assess autistic behaviors. The outcomes of interest were the ASSQ scores, SRS total score, and SRS subscale scores. BLLs were natural log transformed. Negative binomial regression models were used to estimate the association between ASSQ scores and BLLs, and linear regression models were used to estimate the association between BLLs with SRS total and subscale scores. Models were adjusted for sex, fetal and environmental tobacco exposure, paternal and maternal education levels, family income, low birthweight, gestational age, breastfeeding, children's fish intake, and children's blood mercury concentration. There were significant associations between log BLL at 7-8 years of age with ASSQ score (β = 0.151; 95% CI: 0.061, 0.242) and total SRS score (B = 2.489; 95% CI: 1.378, 3.600). Similar associations were not observed for log BLLs measured at 9-10 or 11-12 years of age. In addition, log BLL at 7-8 years of age was significantly associated with the SRS subscores: social awareness (β = 0.438; 95% CI: 0.227, 0.649), social cognition (β = 0.492; 95% CI: 0.209, 0.774), social communication (β = 0.997; 95% CI: 0.121, 0.570), social motivation (β = 0.346; 95% CI: 0.121, 0.570), and social mannerisms (β = 0.217; 95% CI: 0.009, 0.425). Similar results were not observed between log BLLs at 9-10 or 11-12 years of age with the SRS subscores, with the exception of a significant association between log BLL at 11-12 years of age and social mannerisms (β = 0.281; 95% CI: 0.015, 0.547). In sensitivity analyses, significant associations were also observed between log BLL at 7-8 years of age and ASSQ dichotomized into a positive screening result based on a score of 19 as a cutoff (OR = 1.793; 95% CI: 1.010, 3.184) and when using a score of 17 as a cutoff (OR = 1.949; 95% CI: 1.175, 3.232). The authors concluded that there was a positive association between BLLs at 7-8 years of age and more autistic behaviors in children at 11-12 years of age.

Joo et al. (2018) conducted a prospective cohort study to examine gender differences in neurotoxicity due to lead exposure occurring among children from the prenatal period to 5 years of age. Women were recruited from 2006 to 2011 in Korea, and a total of 575 patients were included in the analysis. Blood lead was measured prenatally and when the child was 2, 3, and 5 years of age, and cord blood was measured at birth. Parents' completed the Korean version of the Child Behavior Checklist to assess neurobehavioral development when the child was 5 years old. The generalized linear model and generalized additive model adjusted for maternal age at child birth, parity, maternal educational level, household income, residential area, and breastfeeding. Geometric mean (geometric SD) blood lead levels of the sample were 1.28 (1.48), 1.24 (1.57), 0.90 (1.57), 1.55 (1.49), 1.43 (1.44), and 1.29 (1.38) µg/dL at early pregnancy, late pregnancy, in cord blood, at 2, 3, and 5 years old, respectively. Maternal blood lead levels were higher among those in medium-sized cities compared to those in metropolitan or industrial complexes and those with low household incomes. Blood lead levels were higher among males than females at each time point. A 1 µg/dL increase in late pregnancy blood lead levels was significantly associated with total behavioral problems increasing by 3.00 points (95% CI: 0.56 to 5.45) in males. Among females, increases in age 2 and age 5 blood lead levels of 1 µg/dL were associated with

significant increases in total behavioral scores by 3.82 points (95% CI: 1.25 to 6.39) and 5.72 points (95% CI: 0.44 to 10.99), respectively. The authors concluded that gender differences were present in the effects of lead toxicity on neurobehavioral development, and that males were more susceptible to prenatal exposure and females were more susceptible to postnatal exposure.

Bellinger et al. (1994) examined the association between prenatal and postnatal lead exposure with problem behaviors in children at age 8. Infants born at Brigham and Women's Hospital in Boston, MA between April 1979 and March 1980 were enrolled in the prospective cohort study, and excluded infants who passed away, were adopted or went through custody change, had an unknown location, medical record was not found, or participated in an ongoing study of this birth cohort. Of the 2,657 infants enrolled in the study, 1,782 had completed Teach Report Forms (TRFs) of behavioral outcomes, which included a total problem behavior (TPB) score, internalizing score, externalizing score, and 9 "narrow-band" scales for behaviors. Compared with the group that had missing TRFs, the group that had completed TRFs had lower cord blood lead levels (mean log cord blood lead level of 1.98 vs. 2.02 µg/dL), better neonatal health, had a higher percentage of white children (90.7% vs. 59.7%), and had mothers with more years of education and that were more likely married at the time of delivery. Cord blood lead levels were measured at birth (mean = 6.8, SD = 3.1 µg/dL) and tooth lead levels were measured from a shed tooth at 6 years of age (mean = 3.4, SD = 2.4 µg/g). Both cord blood lead and tooth blood lead were natural log transformed. Multiple linear regression and multiple logistic regression models were used to assess the association between lead levels and behavioral outcomes. Different covariates were used for different exposure/outcome combinations. The models included some subset of the following covariates: prepregnant weight, birthweight, black, cesarean section, paternal education, sex, parents living together during child's first year, aspirin use during week before delivery, maternal education, UTI during pregnancy, current medication use by child, mother married at time of delivery, mother smoking during pregnancy, mother on public assistance at time of delivery, prenatal care begun after first trimester, colic, and sibship size. Cord blood lead level was not significantly associated with TPB, internalizing scores, externalizing scores, or extreme narrow band scores. Tooth blood lead level was significantly associated with TPB (B = 2.06, p=0.0003), internalizing (B = 1.61, p=0.002), and externalizing (B = 1.57, p=0.001) scores. The adjusted odds ratio was not significant for tooth blood lead level and TPB, internalizing, or externalizing scores. The only significant adjusted odds ratio for tooth blood lead level and narrow-band subscales was for "nervous-overactive" behavior (AOR = 1.61, p=0.017). The authors concluded cord blood lead levels were not related to children's risk of behavioral problems, and that a modest association exists between children's postnatal lead exposure (as measured by tooth lead levels) and the risk of behavioral problems.

Longitudinal Studies of Behavior and Lead Exposure

Ethier et al. (2015) conducted a proof-of-concept, prospective cohort study with the primary aim to develop and evaluate a new protocol for effectively measuring visuospatial attention in school-age children with prenatal and current exposure to environmental contaminants. The sample included a total of 27 Intuit children who previously participated in the Cord Blood Monitoring Program in Arctic Quebec, Canada between 1993 and 1998 and were invited to participate in the Nunavik Child Development Study (NCDS) follow-up between 2005 and 2010. The current study took place in 2009 and 2010. The current sample of children had a mean age of 11.2 years (range 8.6 to 12.6), was 66.7% male, had 100% of mothers report smoking during pregnancy, 51.9% of mothers report marijuana use during pregnancy, and 70.4% of mothers report alcohol use during pregnancy. Children's visuospatial attention was tested at the time of the study using a modified version of the Posner paradigm. Additionally, at the time of testing children's blood was measured for levels of mercury, lead, and PCBs. Measurement data

was also available for levels of the same contaminants present in the cord blood at the time of each child's birth, and was used to represent prenatal exposure. The mean cord blood lead level (BLL) of the sample was 5.4 µg/dL (SD = 4.1, range = 1.2 to 17.8), and the mean BLL of the sample at 11 years of age was 3.0 µg/dL (SD = 1.7, range = 1.0 to 7.9). The outcomes of interest obtained from performance on the modified Posner paradigm test included reaction time, omission errors, false alarm errors, accuracy, and validity effect. Multiple regression analyses were used to estimate the associations between contaminant exposures and the attention outcomes described above. BLLs were natural-log transformed. Potential confounders were entered into a model if they had a p-value less than 0.20 and were retained in the model if their addition changed the magnitude of the coefficient for the contaminant of interest by 10% or more. For models that examined cord BLLs and BLLs at age 11, some subset of one or more of the following variables was included: gender, child's age, alcohol during pregnancy, cord DHA, current DHA, cord lead, cord PCBs, and current PCBs. There was a significant association between BLL at age 11 and increased reaction time ($\beta = 0.52$, $p = 0.028$), which the authors reported corresponded to an 81ms increase in reaction time for each one- unit logarithmic increase in BLL based on the raw, unstandardized regression coefficient ($\beta = 81$). In addition, there was a significant association between cord BLL and false alarms ($\beta = 0.42$, $p = 0.015$), which the authors explained corresponded with a 4.4% increase in false alarms for each one-unit logarithmic increase in cord BLL based on the raw, unstandardized regression coefficient ($\beta = 4.4$). The authors concluded that low-level lead exposure during pregnancy was associated with increased false alarm errors representing impulsivity and hyperactivity, while current childhood lead exposure was associated with a slower reaction time reflecting inattention and/or a lack of vigilance.

Wright et al. (2008) investigated potential associations between prenatal and childhood blood lead levels (BLL) with criminal arrests in early adulthood. Pregnant women that attended prenatal clinics located in impoverished neighbourhoods with high concentrations of old, lead-contaminated housing in Cincinnati, OH, were recruited into the prospective birth cohort study between 1979 and 1984. Prenatal BLL was measured, as well as children's BLL which was measured quarterly through 5 years of age and semi-annually from age 5 to age 6.5 years. Criminal arrest data was obtained for a total of 250 participants between 19 and 24 years of age that were included in the present study. The sample was 90% African American, 50% male, 73% from families with scores in the lowest two levels of the Hollingshead SES index, 84% were from households with a single mother, mean age was 22.5 years, and approximately 55% had at least one arrest. The primary outcomes were the individual's total number of criminal arrests beginning at the age of 18, and total number of arrests for violent crimes was also examined. Negative binomial models were used to estimate the association between BLL and arrest rates. Separate models were fit for the three blood exposure variables constructed: prenatal BLL, average childhood BLL (average of each individual's BLL measurements from birth to age 6 or 6.5), and 6 year BLL (measured at either 6 or 6.5 years). Although many covariates including HOME score were considered for inclusion in the models, final models adjusted for maternal IQ, sex, SES (measured with Hollingshead score), and maternal education level. Mean prenatal, average childhood, and six-year BLLs were 8.3, 13.4, and 8.3 µg/dL, respectively. Total arrest rates were significantly greater for each 5 µg/dL increase in prenatal blood lead (RR = 1.40; 95% CI: 1.07 to 1.85) and six-year blood lead (RR = 1.27; 95% CI: 1.03 to 1.57). Arrest rates for violent crimes were significantly greater for each 5 µg/dL increase in average childhood blood lead (RR = 1.30; 95% CI: 1.03 to 1.64) and six-year blood lead (RR = 1.48; 95% CI: 1.15 to 1.89). The authors concluded that prenatal and postnatal lead exposure were associated with higher rates of total arrests and violent arrests and were a risk factor for behaviors leading to criminal arrest.

Choi et al. (2016) reported a study to estimate the incidence rate of ADHD symptoms, and to investigate the blood lead level and parental marital status that influence the development of ADHD symptoms among school-aged children in Korea. Subjects were selected from a previous cohort study, the Children's Health and Environmental Research (CHEER) study which was conducted among elementary school children between 2005 and 2010. Eligible participants were 7 to 9 years of age and did not have ADHD symptoms that were doctor diagnosed or suspected from the baseline survey. Participants were followed for two years. The authors noted that fewer subjects participated from industrial complex residents and more subjects with a family history of psychiatric disorders participated. Parents or caregivers completed a questionnaire, which included an ADHD rating scale and potential risk factors for ADHD, based on the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) adapted into a Korean version. Blood lead level was categorized into quartiles and considered low exposure for ≤2.17 µg/dL and high exposure for >2.17 µg/dL. Analyses were performed via unadjusted and adjusted logistic regression, with the fully adjusted model including residential area, monthly household income, parental marital status, family history of psychiatric disorders, preterm birth and birth weight, child's age, and child's gender. The final cohort in this analysis included 2,159 children who were not suspected of having ADHD during baseline and completed the follow-up survey. The mean blood lead level was 1.56 µg/dL, and was significantly higher among boys ($p<0.001$), children from rural areas ($p<0.001$), children from a home with a low household income ($p<0.001$), and children from single-parent families ($p = 0.025$). ADHD symptoms occurred more commonly among children with blood lead levels >2.17 µg/dL (RR=1.552; 95% CI: 1.002-2.403), as well as among children with a single parent (RR 1.805; 95% CI: 1.002-3.254). Results indicated that family status modified the association between blood lead level and ADHD risk. The authors concluded that lead exposure, including lead exposure at very low concentrations, and adverse family environment were risk factors for the development of ADHD among children.

Summary

Causal associations between exposure to low levels of blood lead (≤ 5 µg/dL) and neurodevelopmental outcomes (whether cognitive or behavioral) in children have not been established given the absence of effective adjustment for well-established causes and risk factors—i.e. confounders—of these neurodevelopmental outcomes in published epidemiological studies, the problems with the reliability of tests for low blood lead levels, and sparse epidemiological data.

Appendix C Part Two: Lead and Behavioral Outcomes (Inattention, Impulsivity, Hyperactivity)

In this part of Appendix C, a systematic assessment is undertaken of the peer-reviewed literature on the relationship between exposure to lead and specific behavioral outcomes of inattention, impulsivity, and hyperactivity.

Background

The purpose of this document is to describe each plaintiffs' expert's opinion on the relationship between lead exposure and behavioral outcomes in children along with the citations they provide to support those opinions.

Additional purposes of this section of Appendix C are to describe and evaluate:

1. the conclusions of recent reports and systematic reviews of the relationship between lead exposure and specific behavioral outcomes in children (inattention, hyperactivity, and impulsivity as well as frank Attention Deficit/Hyperactivity Disorder [ADHD]). These include the NTP Report (2012), the USEPA Report (2014), and the ATSDR Toxicological Profile (2020).

2. the studies and other publications (e.g. systematic reviews and meta-analyses) that were published after these reports (and systematic reviews) appeared, e.g. Donzelli et al. (2019) and He et al. (2019) among others.

Recent Reports and Systematic Reviews of Lead and Behavioral Outcomes

NTP (2012, p. xviii)

> "In children, there is sufficient evidence that blood Pb levels < 5 µg/dL are associated with increased diagnosis of attention-related behavioral problems, greater incidence of attention-related behavioral problems, and decreased cognitive performance…"

The general evidentiary support for the NTP claims regarding attention-related behavioral problems and the diagnosis of attention-related behavioral problems are found in the following table, adapted from NTP (2012, Table 1.2, p. xx):

Population	NTP Conclusion	Principal Health Effects	Blood Pb Evidence	Bone Pb Evidence
Prenatal	Limited Evidence	Increased incidence of attention-related and problem behaviors	Yes <10 µg/dL	No data
Children	Sufficient Evidence	Increased incidence of attention-related and problem behaviors	Yes < 5 µg/dL	Tibia and dentin Pb are associated with attention + behavior

Note that by "Limited Evidence," the NTP writes (2012, p. xv) that "an association is observed between the exposure and health outcome in studies in which chance, bias, and confounding could not be ruled out with reasonable confidence."

By "Sufficient Evidence" the NTP writes (2012, p. xv) that "an association is observed between the exposure and health outcome in studies in which chance, bias, and confounding could be ruled out with reasonable confidence."

USEPA (2014, Table ES-1, p. 1xxxiii)

Re: Externalizing Behaviors: Attention, Impulsivity and Hyperactivity

Causal Relationship (see also Table 4-17, p. 4-414-5)

> "Clear evidence of attention decrements, impulsivity and hyperactivity (assessed using objective neuropsychological tests and parent and teacher ratings) in children 7-17 years and young adults ages 19-20 years. The strongest evidence for blood Pb-associated increases in these behaviors was found in prospective studies examining

> prenatal (maternal or cord), age 3-60 months, age 6 years, or lifetime average (to age 11-13 years) **mean blood Pb levels of 7 to 14 µg/dL** and groups with early childhood (age 30 months) **blood Pb levels >10 µg/dL.** Biological plausibility is provided by animal toxicological studies demonstrating impulsivity or impaired response inhibition with relevant prenatal, lactational, post-lactational and lifetime Pb exposures. Plausible MOAs are demonstrated." *(emphasis added)*

ATSDR (2020, p. 5)

> **"Neurological Effects in Children.** Numerous prospective and large cross-section studies in children provide consistent evidence of decrements in neurological function, including decrements in cognitive function (learning and memory), altered behavior and mood (attention, hyperactivity, impulsivity, irritably, delinquency), and altered neuromotor and neurosensory function (visual-motor integration, dexterity, postural sway, changes in hearing and visual thresholds). These effects have been associated with a PbB range from ≤5 to >50 µg/dL, with numerous studies providing evidence for effects at PbB ≤5 µg/dL."

Note that the ATSDR (2020) report examines only those studies and health effects for lead levels ≤ 10 µg/dL. They write (ATSDR, 2020, p. 133:

> "Due to the enormous size of the database on neurobehavioral effects of Pb, this discussion has been limited to representative and/or major studies published on specific topics crucial to understanding dose-response relationships in the lower exposure ranges (e.g., PbB ≤10 µg/dL). For additional information, the reader is referred to a recent review of this topic (EPA 2014c)."

Here are the studies the ATSDR (2020, Table 2-28, p. 136) cites for behavioral (i.e. altered mood and behavior) outcomes associated with lead levels ≤ 10 µg/dL:

> Arbuckle et al. 2016; Boucher et al. 2012; Braun et al. 2006, 2008; Choi et al. 2016; Dietrich et al. 2001; Froehlich et al. 2009; Fruh et al. 2019; Geier et al. 2018; He et al. 2019; Hong et al. 2015; Huang et al. 2016; Ji et al. 2018; Joo et al. 2017, 2018; Kim et al. 2013a, 2016; Liu et al. 2014a, 2015b; Park et al. 2016; Sioen et al. 2013; Stroustrup et al. 2016; Wang et al. 2008; and Winter and Sampson 2017.

Note that the only mention in the ATSDR report (2020, p. 16) of a method used to assess the scientific evidence is the following:

> "Many specific health effect endpoints have been evaluated in numerous studies. To provide the reader with a **weight-of-evidence** for these endpoints, the profile indicates if results are consistent and corroborated in numerous studies or if results are inconsistent (or mixed)." *(emphasis added)*

Recent Systematic Reviews of Lead Exposure and Inattention, Hyperactivity, ADHD in Children

At least two systematic reviews of the hypothetical relationship between lead exposure and inattention, hyperactivity, ADHD (and similar behavioral problems) have been published recently: He et al. (2019) and Donzelli et al. (2019).

He et al. (2019, abstract), for example, concludes the following:

> "Our findings suggest that low blood lead levels may be associated with ADHD symptoms in children. However, caution is needed when interpreting the results because among-study heterogeneity was in play."

Donzelli et al. (2019, p. 11 of 14) do not make any causal or association claims in their "Conclusions" section. Instead, they simply write that "based on the results of this review, additional data is needed to fully ascertain the nature of the relationship between lead exposure and ADHD." Earlier in the paper, Donzelli et al. (2019, p. 10 of 14) mention that "the results of the present study revealed that in 12 out of 17 studies, a significant association was found between lead exposure and one of the types of ADHD."

It is true that in their abstract, Donzelli et al. (2019) write that "the evidence from the studies allowed us to establish that there is an association between lead and ADHD and that even low levels of lead raise the risk."

It is also true that, on p. 10 of 14, the same authors write that "the current results must be interpreted with caution owing to the presence of a high heterogeneity."

Donzelli et al. (2019) provide a long discussion of the limitations of the studies included in their review:

1. Publication bias
2. The majority of the studies did not validate the diagnosis of ADHD but rather relied upon "data from parents' responses or teachers' responses to behavior checklists that vary from one study to another, and this fact may have produced misdiagnosis or biases."
3. The review also included studies that used urine/teeth lead.
4. Measures of association: there was considerable heterogeneity in the way statistical methods were used.
5. Not all the studies controlled for the same confounding variables and "this could be a source of information bias in this review." Some studies did not consider any confounding variables.

The studies included in the Donzelli et al. (2019) systematic review are as follows:

Chan et al. (2016), Choi et al. (2016), Dikme et al. (2013), Forns et al. (2014), Hong et al. (2015), Huang et al. (2016), Ji et al. (2018), Joo et al. (2017), Kim et al. (2013), Lee et al. (2018), Neugebauer et al. (2014), Park et al. (2016), Sioen et al. (2013), Yang et al. (2018), Yu et al. (2016a), Yu et al. (2016b), Zhang et al. (2015).

More Recent Studies of Pb Exposure and Behavioral Problems in Children and Adolescents

At least two studies have been published on the topic of lead exposure and behavioral problems in children and adolescents since the publication of the reports and reviews mentioned above: Desrochers-Couture et al. (2019) and Lin et al. (2019).

Desrochers-Couture et al. (2019) is the report of an on-going study of Inuit children in Canada who are "more exposed to Pb than children from the general Canadian population (2.34 µg/dL compared to 0.8-0.9 µg/dL, respectively)." In this study, "no significant direct or indirect association was observed between child blood Pb on the adolescent hyperactivity-impulsivity scale" (Desrochers-Couture et al. [2019, p. 8])." Adjustments in the statistical models were as follows: child age, child sex, age of biological mother at delivery, maternal tobacco smoking during pregnancy, birth weight, and cord mercury level (Hg).

Lin et al. (2019) is the report of a study of 164 Chinese children evaluated in a blood lead clinic for symptoms of ADHD. Mean blood lead level for those considered in the "low blood lead" category was 4.3 µg/dL; for those in the "high blood lead" category, the mean blood lead level was 19.6 µg/dL. The authors report that high blood lead levels were not significantly associated with higher hyperactivity/impulsivity levels in the adjusted models. Adjustments in the statistical models were as follows: children's sex, age, birth weight, passive smoking (by parents and others), parity, maternal education, and family yearly income.

Summary

As noted by Lin et al. (2019, p. 165) and taking into account the various publications listed and described above, "there (is) still no consensus on the association between lead exposure and children's hyperactivity/impulsivity." Importantly, even studies of relatively high levels of blood lead do not show an association and, as discussed by He et al. (2019) and by Donzelli et al. (2019)—described above—there is considerable heterogeneity in the reported studies. High heterogeneity is equivalent to inconsistency in results and thus a factor that counts against causation.

Known and Presumed Confounders of the Relationship (if any) Between Low Levels of Lead and Specific Behavioral Outcomes (Inattention, Hyperactivity, and Impulsivity as well as frank ADHD)

As described earlier in this report (see Section 3.2.2) there are at least 19 factors that are known and/or presumed confounders of any relationship between exposure to low levels of lead and the specific behavioral outcomes of inattention, hyperactivity, and impulsivity as well as frank ADHD. These are listed here again along with the systematic review(s) that document these relationships:

- Family history of ADHD and related behavioral symptoms (Azevedo et al. 2018; Van Dongen et al. 2019; Kim et al. 2013; Choi et al. 2016; NIMH website).
- Male gender (Ji et al. 2018)
- Preterm birth (Allotey et al. 2018; Franz et al. 2018; Nielsen et al. 2019)
- Maternal obesity (Cortese and Tessari, 2017)
- Diet/Nutrition (Del-Ponte et al. 2019)
- Maternal smoking (Dong et al. 2018; Huang et al. 2018)

- Air Pollution including particulate matter and NO_2 (Donzelli et al. 2019; Fuertes et al. 2016; Sentis et al. 2017; Thygesen et al. 2020)
- Maternal thyroid dysfunction (Drover et al. 2019; Fetene et al. 2017)
- Acetaminophen use during pregnancy (Gou et al. 2019)
- Vitamin D deficiency (Khoshbakht et al. 2018)
- PBDE exposure (Lan 2017)
- Hypertension of pregnancy (Maher et al. 2017)
- Phthalate exposure (Shoaff et al. 2020)
- Exposure to PCBs (Verner et al. 2015)
- Maternal alcohol consumption (Wetherill et al. 2018)
- Cesarean delivery (Zhang et al. 2019)
- High levels of exposure to other environmental pollutants such as Hg, Cd, Mn, Pb (Kim et al. 2013; Chan et al. 2015; NIMH website)

The importance of these scientific determinations cannot be overstated. Given that these 19 factors (i.e. other than lead itself) can affect (by confounding) any observation regarding the hypothetical relationship between lead exposure and behavioral outcomes (as described, including ADHD) it follows that these factors must be adjusted for in the analyses of lead and behavior. Failure to adjust adds uncertainty to any observation of the lead-behavior relationship. Simply put, studies that do not take these factors into account cannot claim that lead has any effect on the behavioral outcomes of interest here. The basis for this statement is valid scientific methodology. In order to assess the extent to which a factor (like lead) has an effect on an outcome (like behavior), all other known factors must be controlled for in the analysis.

In the section that appears immediately below, I systematically identify the studies of lead exposure and behavioral outcomes (as described above) and provide an assessment of the extent to which the 19 factors listed above were adjusted for in the analyses found in those studies.

Study Characteristics including Adjustment of Confounders of the Relationship Between Lead and Attention Deficit Hyperactivity Disorder and Hyperactivity/Impulsivity Measures (2013-2019)

In this section, the basic design and characteristics of studies on lead exposure and behavioral outcomes, primarily attention deficit hyperactivity disorder and measures of inattention, hyperactivity and impulsivity are described. These studies were identified in several searches:

PMed (9.25.2020) "Metal + Exposure + Children + Behavior + Attention + Epidemiology" (n = 15)
PMed (9.25.2020) "Lead + Exposure + Children + Behavior + Attention + Epidemiology" (n = 19)
PMed (9.25.2020) "ADHD + Children + Environment" (n = 200)

After removal of duplicates, the following studies were identified and will be examined in this section: Braun et al. (2006), Nigg et al. (2008), Ha et al. (2009), Dikme et al. (2013), Kim et al. (2013), Sioen et al. (2013), Forns et al. (2014), Chan et al. (2015), Hong et al. (2015), Neugebauer et al. (2015), Zhang et al. (2015), Arbuckle et al. (2016), Choi et al. (2016), Huang et al. (2016), Park et al. (2016), Yu et al. (2016), Joo et al. (2017), Ji et al. (2018), Lee et al. (2018), Desrochers-Couture et al. (2019), Fruh et al. (2019), and Lin et al. (2019).

Information collected for each study: author & year & location, study design, assessment of diagnosis, pre-post measurement of outcomes? (y/n), mean blood or bone lead level, and confounders adjusted for in the analysis.

Table: Studies of Behavioral Outcomes (Inattention, Hyperactivity, Impulsivity, and ADHD) and Lead Exposure in Children (2006-2019)

Author, Year Location	Design	Assessment ADHD Diagnosis	Pre-Post Measures?	Mean PB level	Confounders Adjusted
Braun et 2006 NHANES (USA)	Cross-Sectional	Parent Recall	No	NR	age, race, sex, prenatal ETS, postnatal ETS, preschool, health insurance, ferritin
Nigg et 2007 Michigan	Case-Control	Testing	No	1.03 µg/dL	age, gender, income
Ha et 2009 South Korea	Cross-Sectional	Parent Tests	No	1.8 µg/dL	age, gender, income, parental hx psych, residence, mercury
Dikme et 2013 Turkey	Cross-Sectional	Clinical Diagnosis	No	1.91 µg/dL	age, sex, mercury
Kim et 2013 Nebraska	Case-Control	Clinical Diagnosis	No	1.29 µg/dL	Age, race, sex, maternal SMK, maternal EtOH, ETS, SES, residence
Sioen et 2013 Belgium	Prospective	Parental Questionnaire	No	1.43 µg/dL	BMI, age, SMK, ETS, education, gender
Forns et 2014 Spain	Prospective	Clinical Diagnosis	No	0.344 µg/dL	age, test quality, sex, maternal IQ, social class, country, maternal mental health, maternal age, other dx
Chan et 2015 USA	Cross-Sectional	Teacher Scale	No	0.5 µg/g (Teeth)	race, sex, education, marital status, SES
Hong et 2015 South Korea	Cross-Sectional	Parents +Teacher Scales	No	1.80 µg/dL	age. sex. region, education, income IQ, Hg, Mn, cotinine, phthalates
Neugebauer et 2015 Germany	Prospective	Child Tests	No	20 µg/L Mother's Blood	gender, age, SMK, EtoH, jaundice, mother's psych, nationality, education
Zhang et 2015 China	Cross-Sectional	Parental Scale	No	7.9 µg/dL	age, sex, ferritin, father's job, e-waste workshop near home
Arbuckle et 2016 Canada	Cross-Sectional	Clinic Questionnaire	No	0.9 µg/dL	Not Described
Choi et 2016 South Korea	Prospective	Parental Report	No	1.44-1.69 µg/dL	age, gender, residence, income, marital status, family hx, preterm birth, birthweight
Huang et 2016 Mexico	Cross-Sectional	Parental Questionnaire	No	3.4 µg/dL	marital status, age, education, SES, SMK

Park et 2016 South Korea	Case-Control	Clinical Diagnosis	No	1.59-1.90 µg/dL	age, birthweight, economic status, parental education, parental smoking
Yu et 2016 Taiwan	Case-Control	Clinical Diagnosis	No	1.57-1.73 µg/dL	age, gender, BMI, maternal age, gestational age, education, SMK, EtOH, family history nervous system disease
Joo et 2017 South Korea	Case Control	Teacher Questionnaire	No	1.49-1.65 µg/dL	education, family history, marital status, teenage mother
Ji et 2018 Boston	Prospective	Clinical Diagnosis	No	2.2 µg/dL	maternal age, race, education, SMK, parity, sex, preterm birth, birthweight
Lee et 2018 Taiwan	Cross-Sectional	Clinical Diagnosis	No	1 µg/dL	none
Fruh et 2019 USA	Prospective	Parent/Teacher Evaluation	No	0.4 µg/dL Maternal	age, sex, IQ, SMK, parity, education, mercury, manganese, ethnicity, income
Lin et 2019 China	Cross-Sectional	Clinical Evaluation	No	4.3-19.6 µg/dL	age, sex, ETS, parity, education, income

In a manner similar to that found in Section of this report, I list here the known confounding factors for a relationship between exposure to lead and the behavioral symptoms of inattention, hyperactivity, and impulsivity (as well as frank ADHD) as adjusted for in the studies described in the table immediately above. Not included here are basic demographic variables, such as age, gender, race, socioeconomic status (including income, education, and residence)

Confounding Factor	# of Studies (%)	Comments
Family history ADHD symptoms	5 (24%)	Ha et (2009), Forns et (2014), Neugebauer et (2015), Choi et (2016), Yu et (2016)
Preterm birth	2 (10%)	Choi et (2016), Yu et (2016)
Maternal obesity	1 (5%)	Sioen et (2013)
Diet/Nutrition	0 (0%)	
Maternal smoking	10 (48%)	Braun et (2006), Kim et (2013), Sioen et (2013), Neugebauer et (2015), Huang et (2016), Park et (2016), Yu et (2016), Ji et (2018), Fruh et (2019), Lin et (2019)
Air Pollution **Particulate Matter** **NO_2**	0 (0%)	
Maternal thyroid dysfunction	0 (0%)	
Acetaminophen use	0 (0%)	
Vitamin D deficiency	0 (0%)	
PBDE exposure	0 (0%)	
Hypertension of pregnancy	0 (0%)	
Phthalate exposure	0 (0%)	
Maternal alcohol	3 (14%)	Kim et (2013), Neugebauer et (2015), Yu et (2016)

Cesarean delivery	0 (0%)	
Mercury exposure	4 (19%)	Ha et (2009), Dikme et (2013), Hong et (2015), Fruh et (2019)
Cadmium exposure	0 (0%)	
Manganese exposure	1 (5%)	Hong et (2015), Fruh et (2019)

The results are clear. The studies of lead exposure and symptoms of ADHD (e.g. inattention, hyperactivity, and impulsivity) have inadequately adjusted for these 19 confounders. It is important to point out that even if one study were to adjust for a particular factor, that adjustment does not apply to any other study. In other words, adjustment (controlling) for confounders is study-specific and does not transfer from one study to another.

Appendix D

Methods of General Causation

Section 5.1 Background on Making Scientific Claims of General Causation

Claims regarding general causation rely primarily on scientific method. Indeed, causal claims rely upon a family of interconnected scientific methods. Although professional expertise, assertions about reasonable degrees of certainty, one's "life experience as a scientist," and judgment are not entirely irrelevant, they are no substitute for method in science. Simply put, scientific methods and their application to the available scientific evidence are by far the most important components of causal inference.

Of importance are methods used to synthesize scientific evidence, i.e. methods used to collect, summarize, and interpret the results of many different studies, often from different scientific sub-disciplines. These methods, which are sometimes called "research synthesis methods" or labeled "weight of evidence" methods (Weed, 2005; Krimsky, 2005), will be discussed below.

Any scientific expert's opinion on what causes a disease (or disorder or condition) in any particular instance—for example, whether exposure to lead causes neurodevelopmental—that is made without a careful description of the methods used to support such an opinion, and without a description of the evidence to which those methods have been applied, an appreciation of the quality of that evidence as well as its capacity to test causal hypotheses, should be regarded as insufficient and of questionable reliability and validity. For definitions of reliability and validity see Appendix B.

Methods of Causal Inference

In current practice, claims about disease causation emerge from the application of causal inference methods to bodies of scientific evidence. These methods include: (1) the general scientific method, (2) epidemiological methods, commonly described in terms of study design types such as cohort studies, case-control studies, and ecologic studies, to name some examples, and (3) methods of research synthesis.

Bodies of Scientific Evidence

By "bodies of scientific evidence" I am referring to systematic collections of scientific studies, found in searches of scientific databases (e.g. PubMed and others such as TOXLINE) and by reviewing the reference lists of published articles. Typically, a body of evidence is comprised of studies published in peer-reviewed scientific journals. There are occasions, however, when additional sources of information are important, such as unpublished studies, abstracts of scientific work, textbooks, and the reports of public health agencies involved in scientific research; these are sometimes referred to as examples of the "grey literature." In this case, for example, reports (including websites) of the National Toxicology Program (NTP), the National Cancer Institute, and the World Health Organization's International Agency for Research on Cancer (IARC) may be discussed and cited.

Scientific studies are often described in terms of their design, with a broad categorization involving either: descriptive studies (such as prevalence surveys, case reports, and case series), analytical studies (such as epidemiologic cohort studies and case-control studies), and experimental studies (such as randomized clinical trials and animal bioassay studies). In addition, any scientific study can be examined in terms of its quality and validity. Quality concerns typically include: the extent to which measurements

of exposure and disease outcome were made accurately and reliably, the appropriateness of the quantitative methods used to analyze the data, and the extent to which those quantitative methods were used appropriately. Validity concerns typically include: the extent to which the results of the study are biased, whether by random error (assessed using tests of statistical significance and/or confidence intervals) or by systematic error. Systematic error (bias) has many different types, including but not limited to: confounding, information bias, and selection bias. Definitions of these biases—all challenges to the internal validity of a study—can be found in textbooks of epidemiology, for example, Aschengrau and Seage (2003a).

Published reviews of the scientific evidence are important sources of information in causal assessments. Not all reviews, however, provide causal conclusions; some, for example, provide recommendations for further research. The quality of any review of the literature—like the quality of an individual scientific study—may affect the validity and reliability of the conclusions and the inferences that should be drawn from that review. The quality of a review is determined, in part, by the extent to which it is systematic, as discussed in more detail below.

Scientific Judgment

Scientific judgment is an important part of the process of causal inference (Weed, 2007). When collecting, summarizing, and interpreting scientific evidence, judgment is applied at many steps. Judgment incorporates scientific reasoning. Judgment is not, therefore, a purely subjective process, although it has a clear and significant subjective component. Judgment is important in causal inference, but it is not, by itself, sufficient for making causal claims. Methods of causal inference and a systematically collected body of evidence are the necessary components of any causal analysis. These methods take precedence over judgment in any causal assessment.

Methods of Causal Inference: General Scientific Method

The general scientific method is the over-arching methodology that must be used in making causal claims about human health and disease. The aim of science (and so its methodology) is explanation, typically causal explanation. Reliable and valid claims regarding causal explanations emerge from the application of methods, including the general scientific method as well as the many more specific methods that have been developed by biomedical scientists over the past half century. Scientific methods provide the objectivity so critical to the validity and reliability of scientific claims about disease causation.

The basic structure of the scientific method includes: a causal hypothesis, observable predictions of that hypothesis, the observations themselves, alternative hypotheses, and tests to distinguish between the causal hypothesis of interest and its alternatives. The process of making causal claims using this basic (although over-simplified) scientific method proceeds as follows: a causal hypothesis is proposed (often, by observing a phenomenon that requires explanation, such as the occurrence of a condition or disease in a single individual (a patient) or a series of patients). Predictions that can be observed in studies are then made. For example, one such prediction is that in a population of similar individuals, some of whom are exposed to a substance (e.g. lead) and others who are not, the incidence of the condition (e.g. negative birth outcome) should be greater in those exposed than in those unexposed with due consideration for alternative explanations (e.g. chance, confounding, and bias).

Epidemiology Studies and the Scientific Method

Comparisons of the type mentioned immediately above are the key contrast found in well-designed analytical epidemiologic studies. Indeed, epidemiology can be described as the basic science of the public's health. Whether designed as prospective or retrospective cohort studies or as case-control studies, epidemiologic studies compare the incidence rate (i.e. occurrence) of a disease or condition (e.g. a birth outcome or a neurodevelopmental outcome) in those exposed to a hypothetical causal factor (e.g. lead) to the incidence rate of the same disease in those not so exposed. In cohort studies, the numerical measure of this hypothetical relationship is called the "relative risk," often abbreviated as "RR." This measure of association is calculated for internal analyses, i.e. analyses comparing the rate of disease in one group in the cohort (e.g. exposed) relative to that of a reference group (i.e. unexposed) from the same cohort. RRs are used to examine whether exposure-response relationships exist between exposure to lead and outcomes such as IQ. Case-control studies (the structure of which can be found in textbooks by Bhopal (2002) and Checkoway et al. (2004)) also involve calculations of a numerical measure that estimates this same relative risk. In those studies, this measure is called the "odds ratio" or "OR."

In contrast, when mortality (i.e. deaths and death rates) are used in the study rather than the incidence of disease, and when the comparison group is an external "standard" population, the numerical measure of this hypothetical relationship is called the "standardized mortality ratio" or "SMR" (Checkoway et al., 2004). SMRs summarize the observed number of deaths in the cohort relative to that which would be expected in a representative general population, such as the US population.

Confidence intervals (95% CI) and p-values can be calculated for both SMRs and RRs; more on these statistical tests will be described below as needed.

Alternative hypotheses are also important in the general scientific method and can affect any study, including but not limited to epidemiology studies. These alternatives include but are not limited to: chance, other factors that may also cause the same outcome (confounders), and bias. Finally, tests are proposed and undertaken to attempt to distinguish the causal hypothesis of interest from its alternatives.

Statistical Hypothesis Testing and the Scientific Method

For example, tests of statistical significance and confidence intervals provide the scientific investigator with information that may allow him to exclude chance as an alternative explanation. If a 95% confidence interval includes the value of 1.0 (for a relative risk estimate), then chance cannot be excluded as an explanation for the results. Similarly, if a test of statistical significance fails to achieve some pre-specified level of significance, such as $p < 0.05$, then chance cannot be excluded as an explanation for the results. Failure to achieve statistical significance introduces uncertainty because an important alternative explanation for the results cannot be dismissed (or excluded). Besides chance, there are many competing explanations for any scientific observation, such as bias and confounding (to be discussed in more detail later in this report). The goal of the scientific method is to discern which explanation is the best (i.e. the most valid and reliable) explanation.

The general scientific method is applied every time a scientific study is undertaken. It is not always the case that investigators explicitly describe each of its steps, although each step is a critical part of the method. It is especially important to recognize the importance of alternative hypotheses when general causation is the issue. For example, if an outcome is observed in a population of individuals who have been exposed to a substance, it is reasonable (and important) from a scientific perspective to ask: what are all the possible causes of this outcome and to what extent have these alternative hypotheses been

excluded? If the outcome occurs in individuals who have never been exposed, then the exposure may not be the cause of the condition in this group (population) of individuals. Only carefully performed scientific studies with control groups can begin to sort out whether the exposure of interest or some other factor causes the outcome in any population. Studies themselves are not sufficient, however. To these studies, collected into a body of evidence, are applied the methods of causal inference and scientific judgment.

Finally, the extent to which an outcome in a single individual was caused by the exposure is not the problem of general causation. That is a very different problem—individual (or specific) causation.

Methods of Causal Inference: Epidemiological Methods

Epidemiological methods involving human subjects (i.e. study participants) are the most important means for identifying and testing hypotheses involving human disease causation. These studies are prioritized as follows from strongest to weakest (USPSTF, 2008):

- Randomized Controlled Trials, Well-conducted Systematic Reviews or Meta- Analyses of Homogeneous RCTs
- Well-designed Controlled Trials without Randomization
- Well-designed Cohort Studies
- Well-designed Case-Control Studies
- Ecologic Studies and Multiple Time Series Studies
- Case Series (i.e. a collection of individual case reports)
- Expert Opinion

By "strongest," I mean the strongest test of a causal hypothesis, i.e. the study design that best distinguishes between the hypothesis of interest and alternative hypotheses. A randomized controlled trial (RCT) is generally considered to be the strongest test of a causal hypothesis in a single study. By "strongest," I do not necessarily mean the study from which the strongest recommendations for some action—e.g. a recommendation to control or regulate exposure to a chemical—should be taken, even though that may, in some specific circumstances, be the case. Any action taken on the basis of the scientific evidence is a complex matter, involving issues well beyond the results of scientific studies and their interpretation. Discussions of specific regulations regarding exposure to lead and their related public health recommendations are not a part of this report unless relevant to the causal question.

By "weakest study" I mean the weakest test of a causal hypothesis. Expert opinion in the absence of evidence and a methodology is an extraordinarily unreliable way to make causal claims. Expert opinion in the absence of evidence and methodology is no test of a causal hypothesis. In this regard, it is also important to point out that case series (and single case studies) are only marginally better than expert opinion. Although they involve actual observations, case series (and single case reports) provide no test of a causal hypothesis (Holland, 1986) and are, therefore, a logical problem to be discussed in more detail below. There are no established methodologies in the published scientific literature on the topic of causal inference that rely upon collections of case reports. There are, however, publications on the role of case reports in causation analyses, primarily in the literature on pharmacovigilance, i.e. the occurrence of adverse drug reactions in the practice of medicine. It is generally held that, even in that situation, making reliable causal claims is severely hampered and not recommended when case reports and case series are used as the primary evidentiary source (Jick, 1977; Venning, 1983). At best, case reports and case series generate (rather than test) causal hypotheses about the relationship between medications and adverse reactions (Sachs and Bortnichak, 1986; Goldman, 1998). In sum, case reports

only provide clues to etiology (Vandenbroucke, 1999). The reason case reports cannot test causal hypotheses will be discussed in the next section.

The Fundamental Problem of Causal Inference

We cannot observe on the same individual both the effect of a cause (e.g. a disease outcome) and what would have occurred had the cause not acted to produce its effect (Holland, 1986). This constraint is sometimes called the "counterfactual" condition. If, for example, an individual is exposed to a chemical and is later diagnosed with an illness or condition, we cannot know whether that same person would have contracted that illness or condition without being exposed.

This fundamental problem of causal inference is the primary reason why randomized clinical trials are considered the gold standard of scientific research in therapeutic and (etiologic) preventive clinical research; randomized trials provide the best approximation to solving this problem by assuming that the individuals who do not receive the intervention—e.g. the placebo controls or controls given some other intervention—are as similar as possible to the individuals who do receive the intervention. Certainly, the control group is not exactly the same as the treated group; but randomization assures that the differences between the two groups are distributed evenly among them.

The use of the randomized clinical trial to "solve" the fundamental problem of causal inference also explains why control groups in any study of human health effects are so important. Epidemiological studies, such as the prospective cohort study and the case-control study, do not have randomly identified controls, but controls in those studies are carefully selected nevertheless, for precisely the same reason they are used in randomized trials: to provide an approximate solution to this fundamental problem. Claims about disease causation from studies that lack control groups are of questionable validity and reliability. Note that simply selecting similar controls is insufficient; potential confounders must be controlled for in the data analysis.

Ecologic studies also do not test causal hypotheses relating to the individuals in a population. Making general causation claims from ecologic study results is inappropriate and has been dubbed the "ecologic fallacy." (McLaren and Hawe, 2005; Porta, 2008, p. 75)

A complete account of the design and analysis of epidemiological studies is beyond the scope of this report. Epidemiology is a mature scientific discipline with graduate degree programs and textbooks describing in detail the methodologies and other issues involved in undertaking an epidemiological study (See, as representative examples, Aschengrau and Seage, 2003, or Bhopal, 2002). The emphasis in this report is on the interpretation of epidemiological evidence for the purpose of making claims about general causation. Given the purposes of this report, I may consider to a limited extent the biological plausibility of the general causal hypothesis at issue in this litigation, e.g. whether a mechanism has been demonstrated that can explain at the biological (cellular) level why exposure to lead causes any number of neurodevelopmental outcomes and has been demonstrated in human populations. Specific methodological issues pertaining to the design and analysis of individual epidemiological studies and toxicological studies (e.g. animal bioassays and studies involving cell cultures) will be discussed in this report as needed.

Methods of Causal Inference: Research Synthesis Methods.

Claims of general causation are made using a body of evidence, that is, a collection of studies. To this body of evidence the methods of causal inference are applied. Several such methods are generally recognized. These are sometimes referred to as research synthesis methods (Weed, 2000a):

1. The systematic narrative review
2. Meta-analysis, including pooled analyses
3. Criteria-based methods of causal inference

These methods are important in a discussion of any hypothetical causal relationship, although there are situations in which meta-analysis should not be performed.

Systematic Narrative Review, or "Evidence-based" Review

This approach was briefly described above. I begin here with a discussion of how to evaluate the quality of review papers.

The quality of review papers published in the scientific literature is an issue of considerable importance, given that it is within these publications that general causation claims most often appear. Put another way, it is inappropriate for general causation claims to appear based on the results of single studies. Assessments of general causation always involve the collection, description, summarization, and interpretation of bodies of scientific evidence comprised of many studies, often from many different disciplines. It follows that the quality of review papers is key. Indeed, a poor quality review cannot be considered "systematic" and its conclusions should be regarded as having questionable validity and reliability.

A good example of the minimum requirements for a scientifically rigorous systematic review can be found in Shea et al., (2007a):

8. a clear description of the purpose of the review,
9. explicit search terms and databases,
10. explicit inclusion and exclusion criteria (for the studies to be reviewed),
11. duplicate data abstraction and a process for resolving disputes between abstractors,
12. explicit consideration of the so-called "grey" literature, i.e. unpublished reports, etc.
13. detailed descriptions (e.g. a table) of the characteristics of the included studies,
14. formal quality assessments of the included studies,
15. appropriate incorporation of the quality assessments in combining results,
16. appropriate methods for combining results of the studies,
17. explicit assessment of publication bias (e.g. in meta-analyses), and
18. an explicit discussion of potential conflicts of interest (e.g. funding sources).

Failure to include all or most of these items in a published review should be interpreted to mean that the review is not systematic. Indeed, these 11 questions form the basis for a valid and reliable tool for evaluating the quality of a review (Shea et al. 2007b; Shea et al. 2009); that tool is called "AMSTAR" and has been extensively used in the scientific literature, including publications involving oncology, the study of cancer (Duncan et al. 2017; Hasan et al. 2017; Kung et al. 2010; Li et al. 2012).

Meta-Analysis

Meta-analysis is a method for combining results across scientific studies. For a discussion of the quality of (and therefore a list of the essential components of) meta-analyses, see Stroup et al., 2000. Under certain conditions, it is possible to average results from several different studies or, alternatively, to pool the data (i.e. using the original individual study participant information from several studies). Pooling, in essence, creates a single study from the data collected from many studies. Only very similar studies—i.e. non-heterogeneous studies by design or by assessing heterogeneity using statistical tests—can be

combined using meta-analytical techniques. The value of meta-analysis is that it increases the statistical power of the effort to identify and quantify statistical associations. As a result, the pooled results (or meta-analyzed results) may be able to detect statistically significant (and typically more precise estimates of) associations that could not be detected in smaller studies. Meta-analysis can also provide an assessment of the consistency of evidence, as described in more detail in the next section (Weed, 2000b; Weed, 2010). Finally, published meta-analyses typically (and appropriately) provide an assessment of publication bias.

Criteria-based Methods

The criteria of causation, often referred to as "Hill's criteria," are traditionally the most important step in causal inference involving both studies of human populations and laboratory-based studies of biological mechanisms (Weed and Gorelic, 1996; Weed, 1997). Typically, these criteria are only applied to a body of evidence once a statistical association has been established, i.e. once several epidemiological studies have revealed increased risks of a disease (or condition) that could be called "statistically significant," most often at the level of ($p < 0.05$), whether evaluated using p-values or confidence intervals. Under these conditions, it is reasonable to conclude that the observed association is not due to chance (Hill, 1965; Hill, 1971; Susser, 1986). These risks are described in terms of incidence rates or mortality rates with corresponding relative risk estimates. The application of the Hill criteria is fully consistent with the well-known phrase that "association is not causation" (Bind, 2019).

It will be helpful and important to describe these so-called criteria, their characteristics, and some of their strengths and limitations, because they have appeared in the scientific literature on causation (more generally) and figure prominently in assessments of general causation in practice.

I use the word "so-called" to describe these criteria (or, alternatively, "factors" or "guidelines") because so many commentators have emphasized the fact that these are not criteria in the classic sense of the word, with one exception: temporality. Note that in this report I will use the term "criteria" to describe these considerations (or factors) in keeping with common usage.

Temporality—the concept that a causal factor must precede its effect—is a true criterion (Susser, 1986; Aschengrau and Seage, 2003; Beaglehole et al., 1993; Gordis, 2000; Goodman and Samet, 2006). In the absence of temporality, i.e. if the presumed effect precedes its hypothetical cause in time, then causation can, for practical purposes, be ruled out. But the existence of the appropriate temporal order—a hypothetical cause followed by its presumed effect—does not establish causation, because many alternative explanations remain to be considered and because causation is not established because one event follows another. Some commentators include the concept of latency under the rubric of temporality. Latency refers to the length of time between the onset of a hypothetical causal exposure and a disease. In short, temporality may support causation when it exists, but there are at least eight other so-called criteria to be considered. The remaining eight are presented in the same order that Austin Bradford Hill, the scientist whose 1965 paper on this topic remains a classic source, discussed them.

Strength of Association: refers to the magnitude of the relative risk estimates observed in the epidemiology studies. Typically, the larger the relative risk (RR), the more likely the observed association is causal (Susser, 1986; Aschengrau and Seage, 2003; Beaglehole et al., 1993; Gordis, 2000; Goodman and Samet, 2006). Relative risk estimates can be obtained from well-designed cohort and case-control studies comparing the incidence of a condition in those exposed to the hypothetical cause to the incidence of the same condition in those unexposed. Small magnitudes of association (sometimes

called “weak” or “modest” associations), e.g. relative risks (RRs) of 2.0 or less, are less likely to represent causal associations. Bias (due especially to uncontrolled and residual confounding) can explain the presence of weak associations.

Consistency of Association: refers to the extent to which scientific results are similar (e.g. in direction and statistical significance) across the entire body of epidemiological evidence. Typically, the more consistent are the results, the more likely the observed association is causal Susser, 1986; Aschengrau and Seage, 2003; Beaglehole et al, 1993; Gordis, 2000; Goodman and Samet, 2006). One of the additional values of meta-analysis is that it provides a quantitative assessment of consistency (Weed, 2000b) through tests of heterogeneity; studies homogeneous enough to be combined are, by definition, consistent. Note, however, that these tests do not provide an assessment of design heterogeneity, i.e. differences in study design prominent enough to preclude combining the results in a meta-analysis.

Biologic Gradient (Exposure-Response): refers to the extent to which the relative risk estimates increase in magnitude as the measure of the exposure increases in the epidemiology studies. Typically, a regularly increasing relationship between exposure and risk estimate is more likely to represent a causal relationship than other patterns (Aschengrau and Seage, 2003; Beaglehole et al., 1993; Gordis, 2000; Goodman and Samet, 2006). Note that a statistically significant exposure-response relationship may still not represent a valid relationship. All that statistical significance does is allow the investigator to exclude chance as an explanation.

Biologic Plausibility: refers to the extent to which a mechanism of action has been proposed, studied, and demonstrated, typically in toxicological and other types of laboratory-based studies. It is generally accepted that as the evidence explaining the mechanism of action for a disease increases, the more likely the association is causal (Weed and Hursting, 1998). A disease mechanism has many features, including but not limited to the many intracellular and extracellular changes that occur from the initiating causal event (e.g. an exposure or some unknown “idiopathic” event) to the subsequent disease event. Indeed, latency (discussed briefly above) can be considered one of many features of a disease mechanism. Assessing biological plausibility also involves distinguishing between what happens in humans and what happens in animals. Although animal testing (also called animal bioassay testing) has been used for many years as a component of assessing biological plausibility, its relevance to human health is under intense scrutiny in the scientific community. The primary concern has always been the extent to which the results of animal testing can be extrapolated to humans. Animal testing typically involves exposing rodents (rats, mice, and hamsters) to excessive doses of the chemical of interest to observe whether these same animals subsequently develop disease (e.g. cancer or other outcomes). The evidentiary concerns about animal testing, however, include the following: (1) that animals are exposed to doses (and durations) that far exceed human exposure conditions, (2) the mechanisms of action in animals are not those found in humans, and (3) the physiology of rodents (and their metabolic pathways) may be different than those in humans. A recent report by the National Research Council (NRC) of the National Academy of Sciences (NAS) calls for a dramatic reduction in the use of animal models to assess human health concerns. The NRC report concludes that animal testing lacks relevance, is cumbersome, and is less accurate than methods based on tests using human cells and cell components (NRC, 2007). In light of this fundamental shift in scientific thinking on the role and relevance of animal testing, any and all results from animal bioassays must be considered relatively unreliable. At best, animal bioassay test results should be considered sources of hypotheses to be tested in studies involving human cells, human cell components, or whole individuals (e.g. in epidemiological studies), where mechanisms and modes of action better represent human (rather than rodent) health.

Specificity: refers to two related ideas. First, it refers to the precision with which the exposure and the outcome can be defined and characterized. For example, studies of exposure to lead and full-scale IQ have more specific measures of exposures and outcomes than, say, studies of exposure to metals and cognition. As Hill noted, the more specific the exposure and disease involved, the stronger the argument in favor of causation. Secondly (and traditionally), specificity also refers to the extent that the disease (outcome) has one or more causes (Gordis, 2000; Aschengrau and Seage, 2003). Neurodevelopmental outcomes have multiple causes. It should be pointed out that not all causes of conditions such as neurodevelopmental outcomes have been identified; put another way, there are unknown (often called "idiopathic") causes of neurodevelopmental outcomes.

Coherence: refers to the extent to which the evidence and hypotheses for the results fit together into a reasonable and well-tested explanation (Susser, 1986). In the classic description of this so-called criterion, coherence was defined as the extent to which the causal hypothesis does not conflict with the available evidence. Coherence can be assessed in terms of the extent to which other causal criteria (or "guidelines") have been met. The more criteria that are satisfied, the more coherent the causal explanation.

Experimentation: refers to the extent to which a randomized clinical trial (e.g. a prevention trial) or an observational intervention study has been undertaken (Goodman and Samet, 2006). This is an uncommon condition to be satisfied using randomized trials in the study of chronic diseases. Note that this criterion does not refer to animal experimentation.

Analogy: the extent to which the purported exposure-disease relationship under consideration is similar (in types and characteristics of evidence) to other relationships, known to be causal or not (Weed, 2018).

Use and Uses of the Causal Criteria

These are widely used criteria with scientific justification; the method has been used for at least fifty years in hundreds, if not thousands, of applications involving many different exposures and many different diseases and conditions. Furthermore, many causes of diseases have been identified using this methodology. Research institutions and governmental regulatory agencies regularly use these criteria. Some examples include: the International Agency for Research on Cancer and the U.S. Environmental Protection Agency (Cogliano et al., 2004; USEPA, 2005).

For the past 45 years, epidemiology textbooks have recommended and discussed the use of these criteria for causal inference. See, for example: MacMahon and Pugh, 1970; Mausner and Bahn, 1974; Kleinbaum et al., 1982; Rothman, 1986; Beaglehole et al., 1993; Weed, 1995; Kelsey et al., 1998; Rothman and Greenland, 1998; Vetter and Matthews, 1999; Gordis, 2000; Rothman, 2002; Aschengrau and Seage, 2003b; Bhopal, 2002; Goodman and Samet, 2006.

Several issues should be understood when these criteria are invoked. In cancer epidemiology, commonly used criteria are: strength, consistency, biological plausibility, and exposure-response (Weed and Gorelic, 1996). Other criteria are also used, however, but in a less regular way. Nevertheless, all of Hill's original criteria are relevant to making causal claims, as Hill and many later commentators have recognized. Indeed, an extensive literature on the use of these criteria exists.

If users of this methodology select some of the criteria but not others, they must provide an explanation and justification for their choices. These explanations and justifications are subject to critical appraisal. It is important that users of this criteria-based methodology explain in detail what evidence is needed to

satisfy each of the criteria they select. For example, if "consistency" is important to their use of the method, they should specify how much, what kinds, and what characteristics of the evidence satisfy that criterion. Without an explanation and justification for their individual choices in the use of the criteria-based method, the causal claims of users are suspect. Furthermore, it is important to examine whether users of the criteria appropriately define the criteria and their accompanying rules of inference, whereby "appropriate" I mean the extent to which the user's definitions conform to accepted definitions in the published scientific literature.

Causal Criteria and the Scientific Method

Finally, it is important to point out that these criteria reflect the application of the general scientific method. In other words, the criteria represent key concerns of the general scientific method and not a substitute for it. Examples abound. The criterion of consistency, for example, reflects the basic scientific principles of replicability and testability. The criterion of strength (of association) reflects the basic scientific concept of critically testing alternative explanations. Experimentation, likewise, reflects the need to test and control for alternative hypotheses. Temporality is a key feature of any causal hypothesis. Specificity reflects the need to test the hypothesis of interest and not some different hypothesis. Biological plausibility incorporates biological explanations (at the cellular level) with those explanations at the level of human populations by examining the extent to which the basic causal hypothesis has been tested in cellular systems. And finally, exposure (sometimes referred to as "dose") response reflects a basic toxicological principle: the greater the dose (or exposure) to a causal agent, the greater the effect. In sum, the criteria and the general scientific method are not only compatible but inseparable.

Regression Analyses

Regression analysis is a statistical technique for evaluating the relationship between two or more variables. A common expression of a simple linear regression equation which illustrates the linear relationship between two variables (x and y) is:

$$y = mx + b,$$

where b represents the y-intercept and the slope of the line is presented by m. The y-intercept (b) also reflects the average value in the population when x is equal to zero. The variable y is commonly referred to as the dependent variable or the response variable. The variable x is commonly referred to as the independent variable or the predictor variable. The objective of simple linear regression is to use data from two variables that are linearly related to generate a linear equation or model to be used in predicting values of the response variable (y) given known values of the predictor value (x). In simple linear regression only one predictor variable is used to predict the response variable.

See Figure 1.

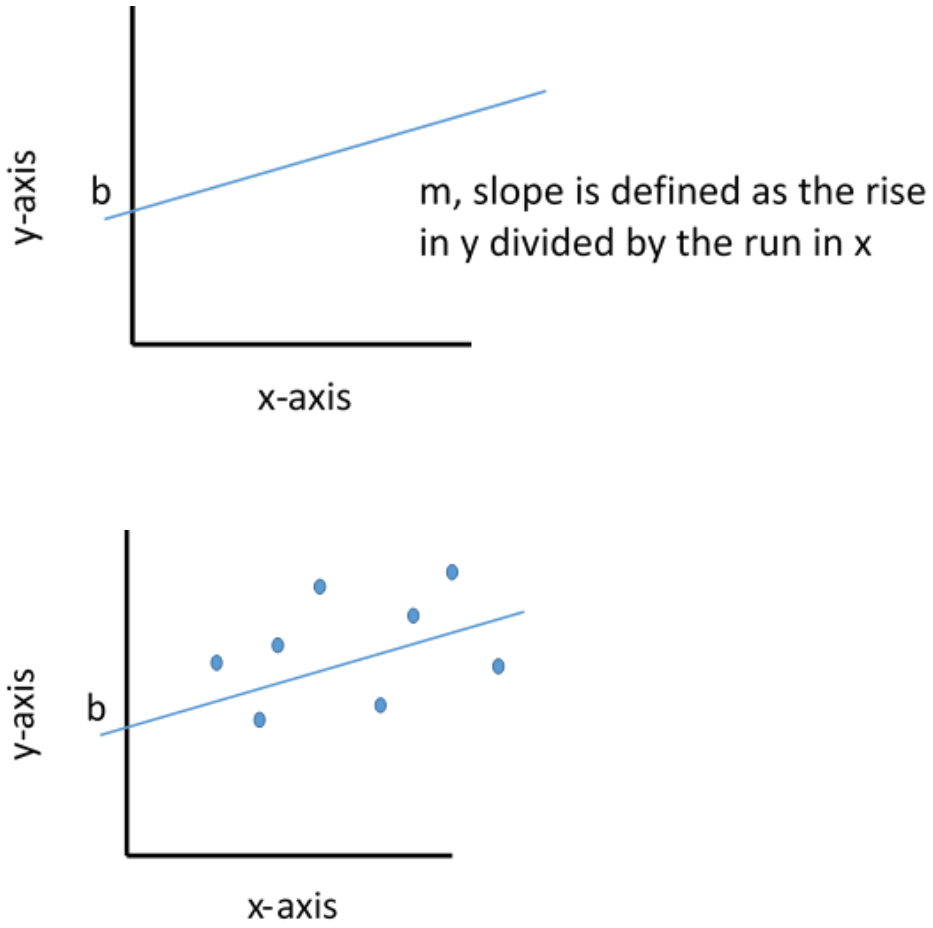


Figure 1. The top graph shows a regression line, with intercept that crosses the y-axis at b. The bottom graph shows the hypothetical data that was used to model the line.

Linear regression usually includes more than one variable, which is known as multiple linear regression. This is where more than one predictor variable is used to estimate the response variable, y. An example of a multiple linear regression equation is:

$$y = m_1x_1 + m_2x_2 + \ldots + m_nx_n + b,$$

where y is the response variable and x_1, x_2, through x_n are the predictor variables. Once again, b represents the y-intercept value. Multiple linear regression is used to account for several factors that could contribute to changes in the response variable, y.

The ability of the regression line to accurately predict the response variable depends on the spread of the data. See Figure 2. When a plot of the data shows data points in a tight line (as the graph on the right), one can have more confidence in their ability to use the regression line to predict the outcome variable y. However, as the data becomes more scattered (as the graph on the left), the ability of the regression line to predict the outcome variable is minimized.

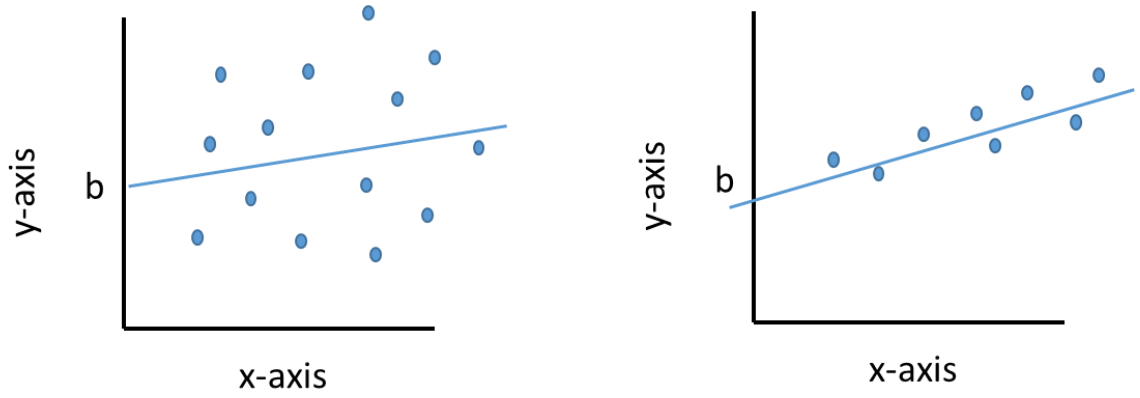


Figure 2. Examples of data point used to generate a line representing the best fit of the data.

One way of measuring the quality of the model is to assess the confidence limits associated with the model. As confidence limits become wider, they indicate less confidence in the model to predict the outcome variable. An example of the confidence limits for the regression lines for the sample data is shown in Figure 3. Confidence limits can also be shown for a specific value of a predictor variable. Figure 4 shows estimated confidence limits for four points. For each point on the x-axis there is a spread of observed y-values and the confidence limits represent the distribution of the data about that given point. The width of the confidence limits are driven by the sample size of data available for a given value of x and the variability of the observed data for a given value of x.

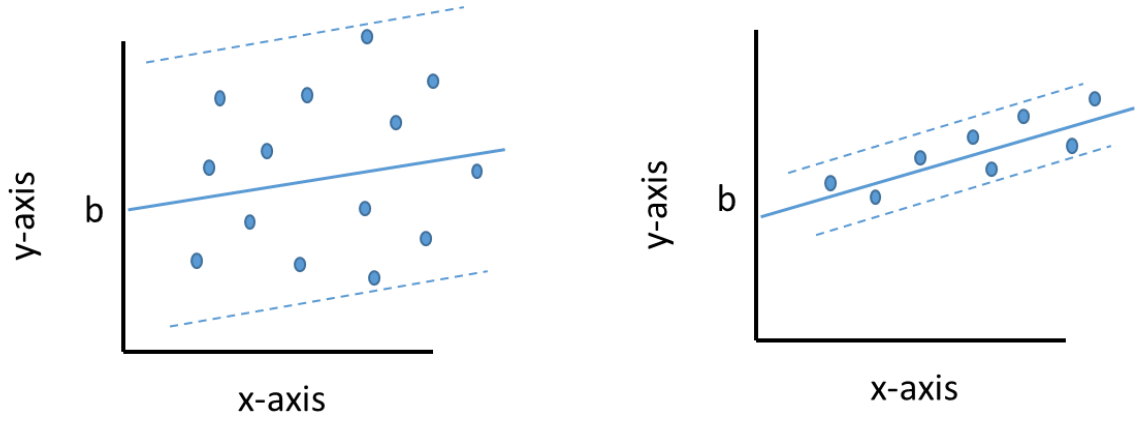


Figure 3. The blue dashed line represent confidence limits on the modeled regression line for the two samples of data.

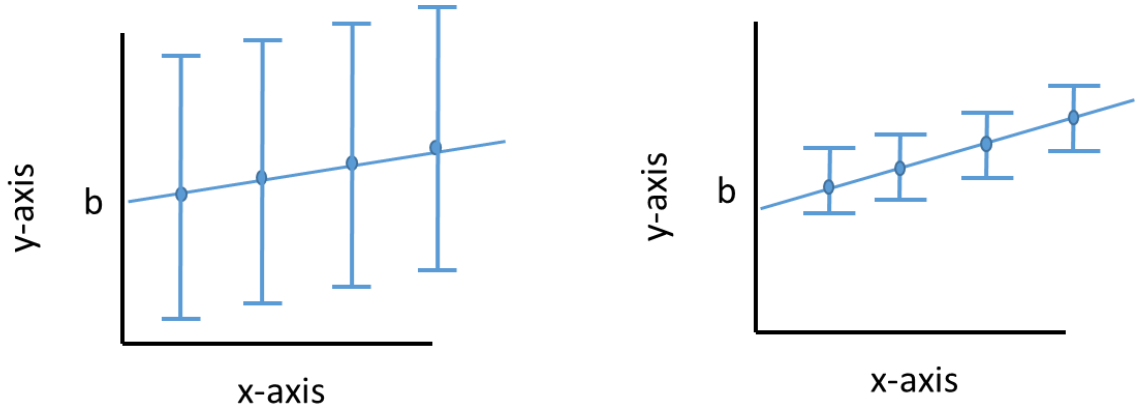


Figure 4. An example of how confidence limits about x reflect the underlying distribution of data for a given predictor value.

It is also important to understand what factors account for the variability in the observed data. For one predictor variable (simple regression) if the fit is not very good and there is a lot of scatter, it is most likely that a single variable is not able to entirely explain the response variable. For multiple linear regression, the variance in the data can be determined for each predictor variable. The higher the measured variance in a particular variable, the more that variable accounts for the scatter observed in the data. In an ideal situation, the sum of the variances of all individual predictor variables would be 100%. This sum of variances is referred to as the coefficient of determination, or R^2, of the regression model. The R^2 of a linear model ranges from 0 to 1, where 1 is a perfect fit. However, it is often the case that there is variance that cannot be fully explained with the model, which is known as unexplained variance.

It is helpful to demonstrate regression using actual response and predictor variables. In this case, the concern is around exposure to lead (measured in children via blood, also known as a blood lead level) and potential impacts on cognitive development (measured as IQ in this example). Figure 5 shows IQ and blood lead level data extracted from Canfield et al. 2003. It is important to note that the graph is depicting lifetime average blood lead. However, the mean peak blood lead levels for the children with complete data for children in this study (n = 154) was above 11.1 ug/dL. The average lifetime blood lead level for these children was 7.7 ug/dL, leaving children with the lowest blood lead levels represented in the graph of Figure 5.

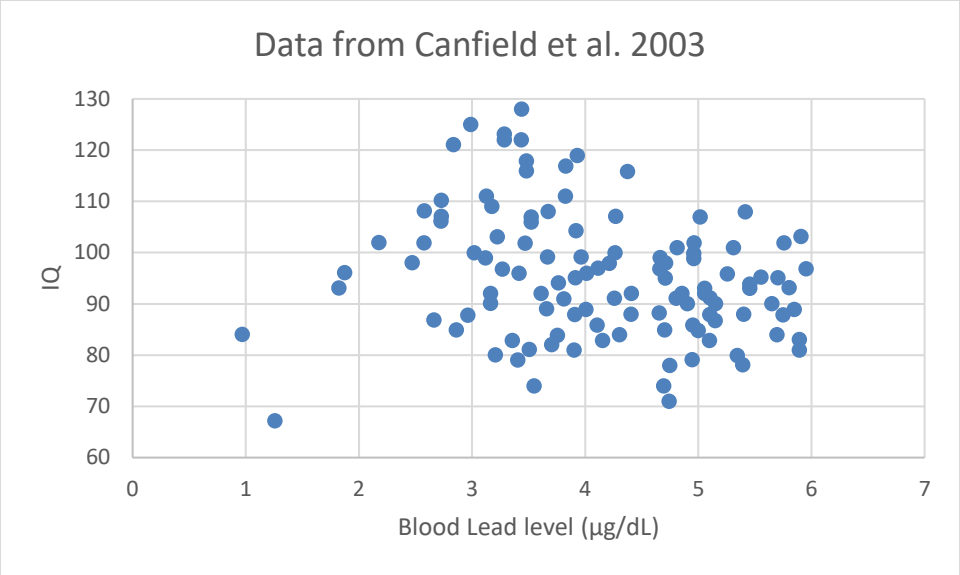


Figure 5. IQ as a function of lifetime average blood lead concentration (original data has BLL as high as 30 µg/dL; focused on low level exposures).

The focus of this example is on low blood lead levels (< 6 µg/dL). Is there an association between blood lead level and observed IQ? A simple linear equation would only consider the variable blood lead. As one can see by looking at the plot, the data points are highly scattered and the child with the lowest average lifetime blood lead level also had the lowest measured IQ. When a linear fit is applied to the data, the relationship between blood lead and IQ is negative, but blood lead level only accounts for 4.79% of the observed variance. This leaves 95.21% of the variance in IQ unexplained.

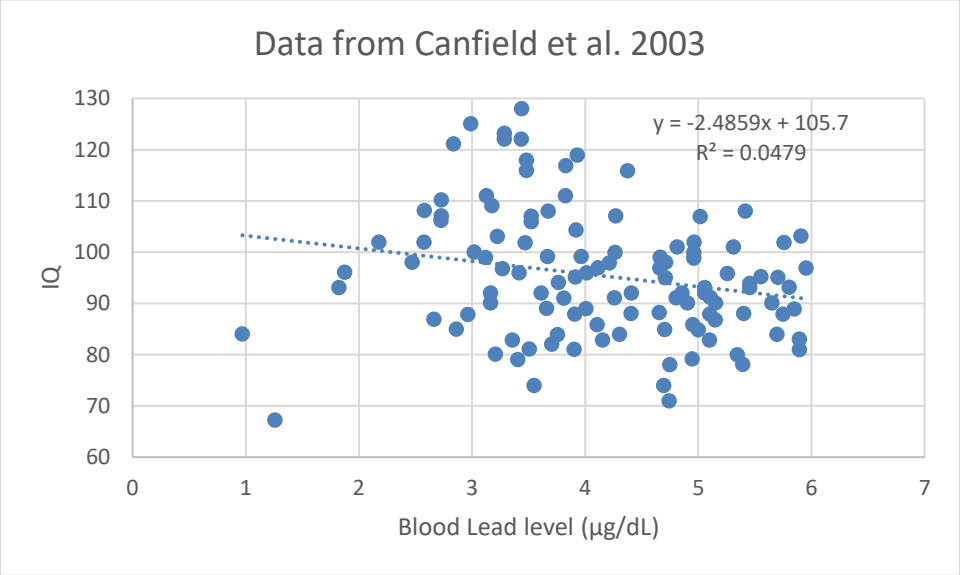


Figure 6. Simple linear regression of blood lead level and IQ with a trendline and R^2 values as estimated in Excel.

Another way to view the data, is to look at groups of children. In Figure 7, children are put into groups based on their average lifetime blood lead level. The data points are as follows: 1 reflects IQ distribution for children with blood lead levels between 0.5 and 1.49 µg/dL; 2 reflects the IQ distribution for children with blood lead levels between 1.5 and 2.49 µg/dL; 3 reflects the IQ distribution for children with blood lead levels between 2.5 and 3.49 µg/dL; 4 reflects the IQ distribution for children with blood lead levels between 3.5 and 4.49 µg/dL; and 5 reflects the IQ distribution for children with blood lead levels between 4.5 and 5.49 µg/dL. As you can see, there are very few data points for average lifetime average blood lead levels near 1 µg/dL. With very few data points (n = 2), the confidence levels are very large and reflect an inability to predict how blood lead levels at such low concentrations may impact IQ. Furthermore, it appears that as blood lead levels increase from 1 to 3 µg/dL, the IQ scores for those children actual increase. However, it is important to note that the error bars for IQ associated with blood lead levels around 2 and 3 µg/dL are broad and indicate that IQ could vary by more than 11 points for children in this category. More children had blood lead levels higher than 3 µg/dL and with each increasing category the sample size increases and the error bars go down. The 95% confidence intervals for children with average lifetime blood lead levels of 4 and 5 µg/dL show a range of 8 and 6 IQ points, respectively.

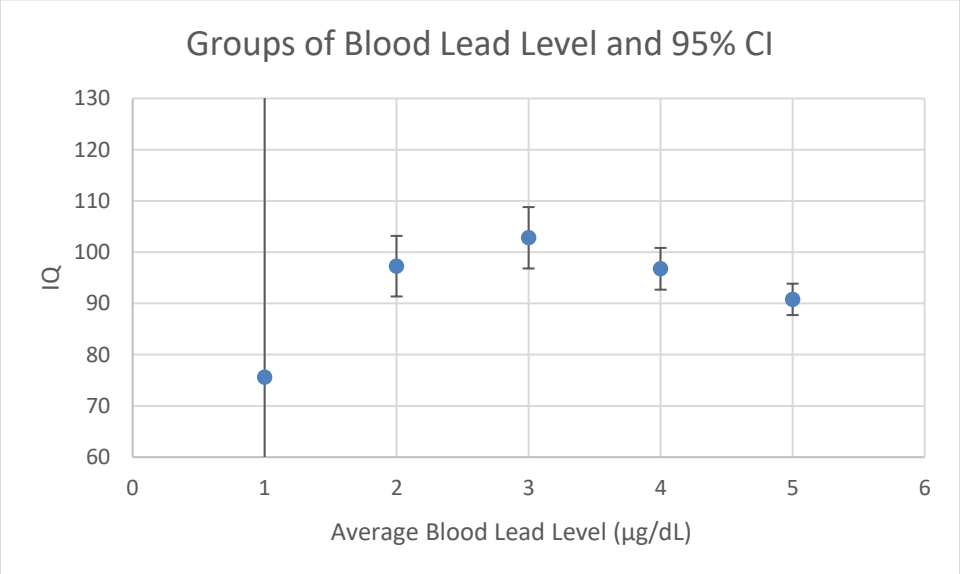


Figure 7. Group distribution of IQ by blood lead level for children with average lifetime blood lead levels less than 6 µg/dL.

All of the specific analyses above, consider the simple model of only blood lead level and IQ in children. However, the literature clearly demonstrates that there are numerous variables to consider when evaluating exposures and cognition. Many authors have evaluated variance contributions of different factors in models, some accounting for as much as 30% of the observed variance. This is why multiple regression analyses are needed in the evaluation of exposure on cognition. Sticking with the Canfield et al. 2003 example, the authors conducted a multiple regression analysis and included several variables in addition to blood lead level: child's sex, birth weight, iron status, maternal measures of IQ, years of education, race, tobacco use during pregnancy, yearly household income, and the total score for the HOME Observation for Measurement of the Environment Inventory. To illustrate the difference including these extra predictor variables have on the final estimate of IQ, the slope of the simple regression line in Figure 6 indicates that for a 1 µg increase in blood lead level, IQ was predicted to

decrease by 2.48 points. However, when the model accounted for other key variables association with childhood cognition, a 1 µg increase in blood lead level (across all measured blood lead levels (0 to 30 µg/dL) was associated with a reduction in IQ of 0.46 points (95% CI: -0.76 to -0.15). For children with average lifetime blood lead levels less than 10 µg/dL a higher reduction in IQ was observed (-1.37 points; 95% CI: -2.56 to -0.17), but note that due to smaller sample sizes the confidence limits get wider. As demonstrated by Crump et al. 2013, in the re-analysis of the Lanphear et al. 2005 pooled analysis, the confidence limits are so large at low exposure levels that conclusions regarding changes in IQ and blood lead levels < 5 µg cannot be drawn from the pooled data.

Appendix E

CURRICULUM VITAE

Douglas L. Weed, M.D., M.P.H., Ph.D.
1302 N. Oak Forest Rd.
Salt Lake City, UT 84103

Phone: 301.980.0197 Fax: 801-359-8939
Email: douglaslweed@aol.com

Education:

1982 – Ph.D., Epidemiology, University of North Carolina

1980 – M.P.H., Epidemiology, University of North Carolina

1977 – M.D., The Ohio State University

1974 – B.Sc., Engineering, *summa cum laude*, The Ohio State University

Experience:

Dr. Weed is an independent scientific consultant. He is a physician-epidemiologist with 37 years of experience in epidemiological research and research training. Dr. Weed is an internationally recognized scholar and educator in causation, causal inference, and the ethics of epidemiology. He has extensive experience in the methods of general causation, cancer causation, systematic reviews, and weight-of-evidence methods. He holds an academic appointment—adjunct full professor—at the University of Utah School of Medicine. He co-chaired the National Academy of Sciences Committee on the 10th anniversary of the U.S. Supreme Court's *Daubert* decision and was a Visiting Scholar at the Federal Judicial Center (Washington, DC). He maintains an active research program in scientific methods, nutritional epidemiology, occupational epidemiology, and the ethics of research. Recent invited lectures include: American Association for the Advancement of Science, at the World Congress of Epidemiology, and at the National Cancer Institute. Dr. Weed is the Reviews Editor for the Journal of the National Cancer Institute and formerly an Associate Editor at the American Journal of Epidemiology.

Dr. Weed is the founder of DLW Consulting Services, LLC. This scientific consulting company provides expertise in disease causation, the methods of causal inference, weight of evidence methods, epidemiological and clinical research methods, and the ethics of epidemiology and

public health. DLW Consulting Services, LLC specializes in providing expert advice and guidance on problems at the interface of science, law, commerce, and public policy. Typical projects include expert testimony and consultation in toxic tort litigation, assessments of health risks from exposure to chemicals, metals, infectious agents, pharmaceuticals, and medical devices, as well as assessments of key methodological and ethical problems facing stakeholders. Examples of such problems include: scientific uncertainty, conflicts of interest, and methods used in legal and regulatory contexts to determine general and specific causation.

Employment:

2008- present	Managing Member, DLW Consulting Services, LLC.
2007-2008	Vice President for Epidemiology and Biostatistics, The Weinberg Group, Washington DC
1990-2007	Chief, Office of Preventive Oncology, National Cancer Institute Director, Cancer Prevention Fellowship Program, Bethesda MD
1982-1989	Senior Staff Fellow, Biometry Branch, National Cancer Institute
1978-1982	Public Health Service Trainee, Department of Epidemiology, University of North Carolina, Chapel Hill, NC.
1978	Research Associate, Environmental Protection Agency, Chapel Hill, NC.
1977	Medical Intern, N. Carolina Memorial Hospital, Chapel Hill, NC.

Professional and Scientific Organizations:

American College of Epidemiology (Fellow)
International Epidemiological Association (Past Member)
Kennedy Institute of Ethics (Member)

Elected Positions:

Board of Directors, American College of Epidemiology, 1998-2001
Executive Committee, Society for Epidemiologic Research, 1996-1999

Editorial Positions:

Associate Editor, Journal of the National Cancer Institute, 1994-present
Reviews Editor, Journal of the National Cancer Institute, 1995-present
Associate Editor, American Journal of Epidemiology, 1997-2013
Editor-in-Chief, NCI Division of Cancer Prevention Newsletter, 1999-2002

Peer Reviewer:

American Family Physician
American Journal of Clinical Nutrition
American Journal of Epidemiology
American Journal of Industrial Medicine
American Journal of Preventive Medicine
American Journal of Public Health
Annals of Epidemiology
Cancer
Cancer Epidemiology
Clinical Trials
Critical Reviews in Toxicology
Diabetes Research and Clinical Practice
Emerging Themes in Epidemiology
Environmental Health Perspectives
Epidemiologic Reviews
Epidemiology
Evidence Based Journal
Food and Chemical Toxicology
International Journal of Epidemiology
International Journal of Health Policy and Management
Journal of the American Medical Association
Journal of Clinical Epidemiology
Journal of Epidemiology and Community Health
Journal of Medical Decision-Making
Journal of the National Cancer Institute
Journal of Toxicology and Environmental Health
Kennedy Institute of Ethics Journal
Nutrition and Cancer
Philosophy and Theory in Biology
Preventive Medicine
Regulatory Toxicology and Pharmacology
Social Science and Medicine
Statistics in Medicine
Theoretical Medicine and Bioethics
Toxicology

Faculty Appointments:

Adjunct Professor, 2014-present
Department of Family and Preventive Medicine
Division of Public Health
School of Medicine
University of Utah

Salt Lake City, UT

Adjunct Professor, 2010 - 2018
Department of Internal Medicine
Division of Epidemiology and Biostatistics
School of Medicine
University of New Mexico
Albuquerque, NM

Visiting Scholar, 2006
Federal Judicial Center
Washington, D.C.

Visiting Fellow, 2001
National Cancer Center
Tokyo, Japan

Visiting Professor (Oncology), 1999
McGill University and University of Montreal
Montreal, Quebec, Canada

Visiting Professor (Epidemiology), 1998
National School of Public Health
Madrid, Spain

Faculty Affiliate, 2001- 2010
Senior Research Fellow, 1995 – 2001
Visiting Fellow, 1994-5
Kennedy Institute of Ethics
Georgetown University, Washington, D.C.

Faculty member, 1994
Society for Epidemiologic Research
Student Workshop on Epidemiologic Methods, Miami, FL

Adjunct Associate Professor, 1994 - 2010
Department of Preventive Medicine and Biometrics
F. Edward Hebert School of Medicine
Uniformed Services University of the Health Sciences
Bethesda, MD

Associate Faculty, 1989 - 2010
Department of Epidemiology
School of Hygiene and Public Health
Johns Hopkins University, Baltimore, MD

Teaching Assistant and Lecturer (Epidemiology), 1979-80
University of North Carolina, Chapel Hill, NC

Honors and Awards:

Engineering Honor Scholar 1971-1974 (each year)
Phi Eta Sigma (freshman academic honorary) 1971
Alpha Epsilon Delta (pre-med academic honorary) 1973
Tau Beta Pi (engineering academic honorary) 1974
Phi Kappa Phi (general academic honorary) 1974
Alpha Omega Alpha (medicine academic honorary) 1977
Honors in Medicine (clinical) 1977
Honors in Obstetrics and Gynecology (clinical) 1977
On-the-Spot Cash Award (NCI): 1999, 2000
Sustained Superior Performance Cash Award (NCI): 1990-1999 (each year)
Distinguished Alumnus: Ohio State Univ. Preventive Medicine 1994
NIH Merit Award 1995
Commencement Speaker: USUHS M.P.H. Graduation 1996
Quality Step Increase (NCI) 1997, 2000
Keynote Speaker: III Congress of Chilean Society of Epidemiology 1997
Keynote Speaker: Spanish Epidemiologic Society 1998
Advances in Oncology Lecture: McGill University Cancer Center 1999
Samuel C. Harvey Lecture: American Association for Cancer Education 1999
Keynote Speaker: Korean Society for Preventive Medicine 1999
Grand Rounds: Ohio State University Cancer Center 1999
Keynote Speaker: Ethics and Research Integrity Day, University of Alberta, 2000
Keynote Speaker: EPA Conference on Environmental Statistics, 2001
J. Walter Juckett Memorial Lecture, Vermont Cancer Center, 2002
Distinguished Leadership Award, NCI Division of Cancer Prevention, 2002
NIH Merit Award, 2004
Keynote Speaker: Great Lakes Cancer Institute Symposium, 2005
Keynote Speaker: Turkish Society of Internal Medicine, 2005

Board and Committee Memberships

Member, Selection Committee (for Medical School Applicants), University of Utah School of Medicine, 2015 - present

Member, Ethics Committee, American College of Epidemiology, 2014 – present

Member, Admissions Committee, University of Utah School of Medicine, 2014 - 2015

Member, Ohio State University College of Public Health Advisory Board
Columbus, Ohio, 2005 – 2013

Member, Commission on Forensic Science and Public Policy, American Judicature Society, 2005 -- 2007

Co-Chair, National Academy of Sciences Committee, 2005 - 2006
"Alternative Models to the *Daubert* Criteria"
Science, Technology, and Law Program, NAS

Chair, Prevention Working Group, 2001-2007
All-Ireland NCI Cancer Consortium
National Cancer Institute (NCI)

Chair, Scientific Education Committee, 1989- 2007
Division of Cancer Prevention, NCI

Chair, Ethics and Standards of Practice Committee, American College of Epidemiology, 1998-2001.

Member, NIH Committee on Continuing Medical Education (CME), 2000-2005

Cancer Advisory Panel, National Center for Alternative and Complementary Medicine, NIH, 1998-2002

World Health Organization Working Group on the Acceptability of Epidemiologic Evidence for Health Impact Assessment, 1999.

National Cancer Institute Cancer Training Advisory Committee, 1997-9.

Member, Advisory Committee for the National Center for Training in Cancer Prevention and Control, Centers for Disease Control and Prevention, 1995-7.

NIH Epidemiology and Clinical Trials Interest Group, 1985-2000.

NIH Committee on Generic Postdoctoral Research Training, 1994.

NCI Committee on Employee Mentoring, 1994.

Program Planning Committee, American Society of Preventive Oncology, 1991-1993.

American Cancer Society Task Force on Preventive Medicine Training, 1993.

NIH Planning Committee for the Alternative Medicine Technology Assessment Meetings, 1993.

ICCCR International Conference on Cancer Prevention. Bethesda, Maryland, February, 1991. See also: Monographs of the Journal of the National Cancer Institute. NIH Publication 91-3227, p.167, 1992.

American Society of Preventive Oncology Annual Meeting Symposium on Quality of Prevention Research. 1991.

Leader, Roundtable Discussion on Causal Inference. Society for Epidemiologic Research Annual Meeting, 1994.

Panel on Philosophy of Science in Epidemiology. Third Brazilian Congress of Epidemiology, Salvador, Bahia, Brazil, 1995.

Leader, Roundtable Discussion on Methods and Morals in Epidemiology. Society for Epidemiologic Research Annual Meeting, 1995.

NCI Roundtable Discussion on Clinical Trials Auditing, 1995.

Leader, Roundtable Discussion on Preventing Scientific Misconduct. Society for Epidemiologic Research Annual Meeting, 1996.

Education Review Committee, U.T. M.D. Anderson Cancer Center, Cancer Prevention and Education Program, 1996-1998.

Member, Ethics and Standards of Practice Committee, American College of Epidemiology, 1996-1998.

Research Interests:

Disease causation, cancer epidemiology, prevention and control, causal and preventive inference, research synthesis methods (evidentiary methods, meta-analysis, systematic reviews, inferential methods, ethical decision-making methods), philosophy of public health, ethics of biomedical research, professional ethics, medical humanities, research training, science and the law.

Recent Lectures and Invited Seminars

"Causality and Epidemiology." DRI Seminar on Toxic Torts and Environmental Law, Phoenix, AZ, February, 2020.

"Best Practices: Causal Inference." University of Indiana School of Public Health, Bloomington, IN, July 2019.

"Case-Based Causality: Application of Artificial Intelligence to Epidemiology and Public Health." George Washington University, Washington, DC. March 5, 2019.

"Narrative and Systematic Reviews: Basics, Quality, and an Innovation." National Cancer Institute, November 14, 2018.

"Case-Based Causality." Queens University, Belfast, Northern Ireland, September 26, 2018.

"Science, Ethics, and Policy: Causation and Conflict." National Cancer Institute, September 6, 2018.

"After the NCI: Methods, Mountains, and Mischief." National Cancer Institute, June 7, 2016.

"Risk of Bias and Causality Assessments." United States Environmental Protection Agency, Crystal City, VA, December 16, 2015.

"Best Practices: Interpreting Observational Studies." University of Alabama, Birmingham. July 20, 2015 and July 25, 2016.

"But are you a good epidemiologist?" Society for Epidemiologic Research Graduate Student Workshop. Denver, CO. June 16, 2015.

"Systematic review and meta-analysis of sugar-sweetened beverages and type 2 diabetes." 33rd International Symposium on Diabetes and Nutrition. Toronto, Canada. June 11, 2015.

"Comments on Scientific Question #1 on Butyl Benzyl Phthalate (BBP) and USEPA IRIS Preliminary Materials." United States Environmental Protection Agency (USEPA). Integrated Risk Information System (IRIS) Bimonthly Public Science Meeting. Crystal City, VA. February 26, 2015.

"Meta-Analyses of Observational Studies." International Life Sciences (ILSI) Annual Conference. Phoenix, AZ, January 19, 2015.

"Causality in Public Health and Preventive Medicine." Department of Family and Preventive Medicine, University of Utah, Salt Lake City, UT, April 18, 2014.

"On the Utility of Criteria-Based Methods of Causal Inference." Society for Risk Analysis. Baltimore, MD, December 9, 2013.

"What Causes Cancer?" Huntsman Cancer Institute. Salt Lake City, UT, November 13, 2013.

"Does Red Meat Cause Colon Cancer?" Center for Advanced Study at the Norwegian Academy of Science and Letters. Oslo, Norway, November 6, 2013.

"Interpreting Scientific Evidence for Cancer Prevention." National Cancer Institute Summer Curriculum on Cancer Prevention and Control. Rockville, MD, July 10, 2014; July 8, 2015; July 8, 2016.

"Conflicts of Interest." University of California, Berkeley. Epidemiology Doctoral Seminar. Berkeley, CA, April 10, 2013.

"On the Utility of Criteria-Based Methods of Causal Inference." International Society for Environmental Epidemiology. Columbia, SC, August 30, 2012.

"How do we make causal conclusions from the 'totality of the evidence' objective and observable?" Conference on "Scientific Approaches to Strengthening Research Integrity in Nutrition and Energetics" sponsored by the University of Alabama, Birmingham. New Paltz, NY, August 2012.

"Standards of Reporting Dietary Supplements Research Studies." National Institutes of Health Office of Dietary Supplements Research Practicum. Bethesda, MD, June 2012.

"Quality of peer-reviewed published reviews: a case study of sugar-sweetened beverages and health outcomes." Institute of Medicine Food Forum. Washington, DC, September 2011.

"Registration of Epidemiological Studies" Pre-Conference Course on Epidemiological Methods, International Epidemiological Association World Congress of Epidemiology. Edinburgh, Scotland, August 2011.

"Comments on Weight of Evidence" AAAS Conference, Washington DC, February 2011.

"The Professional Responsibilities of Epidemiologists." University of California, Berkeley. March, 2010.

"Causal Inference in Cancer Epidemiology." University of California, Berkeley. March, 2010.

"Uncertainty and Weight of Evidence in Risk Assessment." ICNIRP Workshop: Evaluation and Communication of Scientific Evidence and Uncertainty - Towards a Consistent Terminology in Non-ionizing Radiation. Salzburg, Austria, November, 2009.

"Meta-analysis and causal inference: a case study of benzene and non-Hodgkin's lymphoma." Benzene09, Munich, Germany, September, 2009.

"Biological Mechanism and Causal Inference." Institute of Medicine, Washington DC, June 2009.

"A Method for Individual Causation." University of North Carolina, Chapel Hill, NC, May 2008, the American Association of Law Schools Conference on Evidence, Cleveland, Ohio, June 2008, and at Michigan State University, East Lansing, MI, October, 2009.

"Weight of Evidence and Uncertainty Assessments" DIA/FDA Workshop on Risks and Benefits, Bethesda, MD, November 2009 and ICNIRP/WHO Workshop on Risk Assessment and Terminology, Salzburg, Austria, November 2009.

"Cases and Causes" AstraZeneca Wilmington DE, November 2007, and Amgen Inc. Thousand Oaks, CA, March 2008.

"Why should epidemiology bridge the science/law "cultural chasm"? North American Epidemiology Congress plenary session, Seattle, Washington, June 2006.

"Rethinking Epidemiology" Imperial College (London), Division of Epidemiology, London, England, May 2006.

"Weight of Evidence and General Causation" Science for Judges Program, Brooklyn Law School, Brooklyn, NY, March 2006.

"Weight of Evidence: a Review of Concept and Methods." Society for Risk Analysis, Orlando, Florida, December 2005.

"The Future of Cancer Prevention" Keynote Address. Symposium, San Antonio Cancer Institute, San Antonio, Texas, November 2004; and Special Lecture at the 250th Anniversary of the Meath Hospital, Dublin, Ireland, October 2003.

"The End of Epidemiology" Columbia University, Department of Epidemiology, May 2004, University of New Mexico, May 2005 and 2010, Imperial College (London) Department of Epidemiology and Public Health, December 2005.

"Cancer Prevention in the USA" Xi'an Cancer Hospital, Xi'an, China; CICAMS Cancer Hospital, Beijing, China, October 2004.

"Biologic plausibility and other challenges to the primary prevention of cancer." American College of Preventive Medicine, Washington DC, February 2005.

"The Future of Cancer Epidemiology." Michigan State University Department of Epidemiology, East Lansing, MI, April 2005, and the University of New Mexico, Department of Family and Community Medicine, Albuquerque, NM, May 2005.

Advisory Positions

American Health Foundation, 1998-1999.
Australian Cancer Society, 1999.
Health and Environmental Sciences Institute, 2004 – 2005.
International Life Sciences Institute, 2000 – 2003.
World Health Organization, 1999, 2001.
Mead Johnson Nutrition Safety Advisory Panel, 2012 – 2015.
National Science Teachers Association, 2002-2017, 2019.
Brooklyn Law School, 2003, 2006.

Dissertation and Thesis Committees

Vrije University, Brussels, Belgium (Guido Goelen, M.D., Ph.D), 1999-2001

Community Service

Board Member, Local Homeowner's Association, 2017-present
Volunteer, Best Friends Utah, 2018-present

BIBLIOGRAPHY

PUBLICATIONS

Weed DL. 2018. Analogy in causal inference: rethinking Austin Bradford Hill's neglected consideration. Ann Epidemiol 28:343-346.

Weed DL. 2018. The need for systematic reviews in oncology. JNCI J Natl Cancer Inst 110:812-814.

Weed DL. 2016. Causal inference in epidemiology: potential outcomes, pluralism, and peer review. Int J Epidemiol 45:1838-1840.

Alexander DD, Miller PE, Vargas AJ, Weed DL. 2016. Meta-analysis of egg consumption and risk of coronary heart disease and stroke. J Am Coll Nutr 35:204-16.

Acquavella J, Garabrant D, Marsh G, Sorahan T, Weed DL. 2016. Glyphosate epidemiology expert panel review evaluating the relationship between glyphosate exposure and non-Hodgkins lymphoma or multiple myeloma. Crit Rev Toxicol 46(Suppl 1):28-43.

Williams G, Aardema M, Acquavella J, Berry C, Brusick D, Burns M, Camargo J, Garabrant D, Greim H, Kier L, Kirkland D, Marsh G, Solomon K, Sorahan T, Roberts A, Weed DL. 2016. A review of the carcinogenic potential of glyphosate by four independent panels and comparison to the IARC assessment. Crit Rev Toxicol 46(Suppl 1):3-20.

Alexander DD, Weed DL. 2016. On the need for improved methodological quality of published reviews. Am J Clin Nutr 103:683-4.

Alexander DD, Weed DL, Miller P, Mohamed M. 2015. Red meat and colorectal cancer: a quantitative update on the state of the epidemiologic science. J Am Coll Nutr 34:521-43.

Alexander DD, Bassett JK, Weed DL, Barrett EC, Watson H, Harris W. 2015. Meta-analysis of long-chain omega-3 polyunsaturated fatty acids (LCω-3PUFA) and prostate cancer. Nutrition and Cancer March 31:1-12 [epub ahead of print].

Althuis MD, Weed DL, Frankenfeld CL. 2014. Evidence-based mapping of design heterogeneity prior to meta-analysis: a systematic review and evidence synthesis. Systematic Reviews 3:80.

Miller PE, Alexander DD, Weed DL. 2014. Uncertainty of results in nutritional epidemiology. Nutrition Today 49:147-5.

Oostindjer M, Alexander J, Amdam GV, Andersen G, Bryan NS, Chen D, Corpet DE, De Smet S, Dragsted LO, Haug A, Karlsson AH, Kleter G, de Kok TM, Kulseng B, Milkowski AL, Martin RJ, Pajari A-M, Paulsen JE, Pickova J, Rudi K, Sødring M, Weed DL, Egelandsdal B.

2014. The role of red and processed meat in colorectal cancer development: a perspective. Meat Science 97:583-96.

Alexander DD, Weed DL, Chang ET, et al. 2013. A systematic review of multivitamin-multimineral use and cardiovascular disease and cancer incidence and total mortality. J Amer Coll Nutr 32:339-54.

Althuis MD, Weed DL. 2013. Evidence mapping: Methodological foundations and application to intervention and observational research on sugar sweetened beverages and health outcomes. Am J Clin Nutr . doi: 10.3945/ajcn.113.058917.

Weed DL. 2013. The quality of nutrition and cancer reviews: a systematic assessment. Crit Rev Food Sci Nutrition 53:276-86.

Alexander DD, Weed DL, Mink PJ, et al. 2012. A weight-of-evidence review of epidemiologic studies of colorectal cancer in pesticide applicators. Int Arch Occup Environ Health 85:715-45.

Weed DL, Althuis MD, Mink PJ. 2011. Quality of reviews on sugar-sweetened beverages and health outcomes. Am J Clin Nutr 94:1340-7.

Alexander DD, Weed DL, Cushing CA, Lowe KA. 2011. Meta-analysis of prospective studies of red meat consumption and colorectal cancer. Eur J Cancer Prevention 20:293-307.

Navia JL, Byers T, Djordjevic D, Hentges E, King J, Klurfeld D, Llewellyn C, Milner J, Skrypec D, Weed DL. 2010. Integrating the totality of food and nutrition evidence for public health decision making and communication. Crit Rev Food Sci Nutrition 50:1-8.

Schreider J, Barrow C, Birchfield N, Dearfield K, Devlin D, Henry S, Kramer M, Schappelle S, Solomon K, Weed DL, Embry MR. 2010. Enhancing the credibility of decisions based on scientific conclusions: transparency is imperative. Tox Sci doi: 10.1093/toxsci/kfq102.

Weed DL. 2010. Meta-analysis and causal inference: a case study of benzene and non-Hodgkin's lymphoma. Ann Epidemiol 20:347-355.

Weed DL. 2009. Conflicts of Interest. J Epidemiol Commun Health 63:601-602.

Weed DL. 2008. Truth, Epidemiology, and General Causation. Brooklyn Law Review 73:651-665.

Griego FY, Bogen KT, Price PS, Weed DL. 2008. Exposure, epidemiology and human cancer incidence of naphthalene. Reg Tox Pharmacol 51:22-6.

Weed DL. 2007. The nature and necessity of scientific judgment. J Law Policy 15:135-164.

Collins JJ, Bukowski JA, Weed DL, Brent RL, Klein P, Boerstoel-Steefland M, Sprafka JM, Williams AL, Holsapple MP. 2007. Evaluating emerging issues in epidemiology. Regul Toxicol Pharmacol 48(3):296-307.

Dores G, Chang S, Berger VW, Perkins S, Hursting SD, and Weed DL. 2006. Evaluating research training outcomes: experience from the Cancer Prevention Fellowship Program at the National Cancer Institute. Acad Med. 81(6):535-541.

Weed DL. 2006. Commentary: Rethinking epidemiology. Int J Epidemiol. 35(3):583-586.

Parascandola M, Weed DL, and Dasgupta A. 2006. The Surgeon General's Reports on Smoking and Cancer: a historical investigation of the practice of causal inference. Emerg Themes Epidemiol. 3:1-11.

Weed DL. 2006. Evidence synthesis and general causation: key methods and an assessment of reliability. Drake Law Review 54:639-650.

Weed DL. 2006. Vision, values, and verisimilitude: a response. Risk Anal. 26:577.

Weed DL. 2005. Hail the heroes who brave front line. The Times Higher Education Supplement (London) No. 1722:14.

Weed DL. 2005. Weight of evidence: a review of concepts and methods. Risk Anal. 25:1545-1557.

Chang, S., Hursting, S., Perkins, S., Dores, G., and Weed, D.L. 2005. Adapting postdoctoral training to interdisciplinary science in the 21st century: the Cancer Prevention Fellowship Program at the National Cancer Institute. Acad Med. 80(3):261-265.

Resnik, D.B., Kopelman, L.M., and Weed, D.L. 2004. What is the role of the precautionary principle in bioethics and the philosophy of medicine? J Med Philos. 29(3):255-258.

McKeown, R.E. and Weed, D.L. 2004. Ethical choices in survey research. Soz Praventivmed. 49:67-68.

Weed, D.L. and Dores, G. 2004. Physicians as citizens (letter). JAMA 291:2076.

Kopelman, L.M., Resnick, D., and Weed, D.L. 2004. What is the role of the precautionary principle in the philosophy of medicine and bioethics? J Med Philos. 29:255-258.

Weed, D.L. 2004. Precaution, prevention, and public health ethics. J Med Philos. 29:313-332.

Weed, D.L. and McKeown, R.E. 2003. Science and social responsibility in public health. Environ Health Perspect. 111:1804-1808.

Weed, D.L. 2003. Methodologic implications of the precautionary principle: causal criteria. Eur J Oncology. 2:103-108.

Weed, D.L. 2003. Causation: an epidemiological perspective. J Law Policy 43:43-53.

Weed, D.L. and McKeown, R.E. 2003. Science, ethics and professional public health practice. J Epidemiol Commun. Health 57:4-5.

McKeown, R.E., Weed, D.L., Kahn, J., and Stoto, M. 2003. American College of Epidemiology Ethics Guidelines: Foundations and Dissemination. Sci Eng. Ethics 9(2):207-214.

Weed, D.L and Mink, P. 2002. Roles and responsibilities of epidemiologists. Ann Epidemiol. 12:67-72.

Weed, D.L. 2002. Environmental epidemiology: basics and proof of cause-effect. Toxicology 181-182:399-403.

McKeown, R.E. and Weed, D.L. 2002. Glossary of ethics in epidemiology and public health: II. Applied terms. J Epidemiol Commun. Health 56:739-741.

Weed, D.L. 2002. Cancer prevention and the All-Ireland NCI Cancer Consortium. Promoting Health-The Journal of Health Promotion for Northern Ireland 18:26-27.

Weed, D.L. 2001. Methods in epidemiology and public health: does practice match theory? J Epidemiol Commun. Health 55:104-110.

Piniewski-Bond, J.F., Buck, G.M., Horowitz, R.S., Schuster, J.H.R., Weed, D.L., and Weiner, J.M. 2001. Comparison of information processing technologies. J Am Med Informatics Assoc. 8:174-184.

Weed, D.L. 2001. A radical future for public health. Int J Epidemiol. 30:440-441.

Merrill, R.M. and Weed, D.L. 2001. Measuring public health burden of cancer through lifetime and age-conditional risk estimates. Ann Epidemiol. 11:547-553.

Burke, W., Coughlin, S.S., Lee, N.C., Weed, D.L., and Khoury, M. 2001. Application of population screening principles to genetic screening for adult-onset conditions. Genet Test. 5:201-211.

Parascandola, M. and Weed, D.L. 2001. Causation in epidemiology. J Epidemiol Commun. Health 55:905-912.

Weed, D.L. and McKeown, R.E. 2001. Glossary of ethics in epidemiology and public health: I. Technical terms. J Epidemiol Commun. Health 55:855-857.

Weed, D.L. 2001. Theory and practice in epidemiology. Ann NY Acad Sci. 954:52-62.

Stoto, M.A., Hermalin, A.I., Li, R., Martin, L., Wallace, R.B., and Weed, D.L. 2001. Advocacy in epidemiology and demography. Ann NY Acad Sci. 954:76-87.

Weed, D.L. 2000. History of Cancer Prevention: Pioneers of progress: Major Greenwood, Austin Bradford Hill, and the development of the randomized clinical trial (1900-1950). PreventionPost 2:2-3.

Weed, D.L. 2000. The rise and fall of the clinical trial paradigm. PreventionPost 2:1:16.

Weed, D.L. 2000. Interpreting epidemiologic evidence: how meta-analysis and causal inference methods are related. Int J Epidemiol. 29:387-390.

Connor, R.J., Boer, R., Prorok, P.C., and Weed, D.L. 2000. An investigation of design and bias issues in case-control studies of cancer screening using microsimulation. Am J Epidemiol. 151:991-998.

Weed, D.L. 2000. Epidemiologic evidence and causal inference. Hematology/Oncology Clin N America 14:797-807.

Weed, D.L. 2000. History of Cancer Prevention: Joseph Cullen: Champion of Cancer Prevention and Control. PreventionPost 2:2:10.

Weed, D.L. 2000. Heroes and champions. PreventionPost 2:2:12.

Weed, D.L. 1999. Towards a philosophy of public health. J Epidemiol Commun. Health 53:99-104.

Potischman, N. and Weed, D.L. 1999. Causal criteria in nutritional epidemiology. Am J Clin. Nutrition 69(Suppl):1309S-1314S.

Weed, D.L. and Coughlin, S.S. 1999. New ethics guidelines for epidemiology: background and rationale. Ann Epidemiol. 9:277-280.

Weed, D.L. 1999. Book review of: *Philosophy in Epidemiology and Public Health*. Epidemiol. Monitor.

Ratnasinghe, D.D., Weed, D.L., and Shankar, S. 1999. Cancer knowledge and misconceptions among Hispanic El Salvadorian men in the Washington DC area. J Immigrant Health 1:207-213.

Weed, D.L. and Hursting, S.D. 1999. The authors reply (letter). Am J Epidemiol. 150:218-219.

Weed, D.L. 1999. History of Cancer Prevention: Cancer prevention in the "Roaring Twenties." PreventionPost 1:6.

Weed, D.L. 1999. The DCP Newsletter team (editorial). PreventionPost 1:12.

Weed, D.L. 1999. Higher standards for epidemiologic studies-replication prior to publication? (letter). JAMA 282:937.

Weed, D.L. 1998. Preventing scientific misconduct. Am J Public Health 88:125-129.

Weed, D.L. 1998. Beyond black box epidemiology. Am J Public Health 88:12-14.

Breslow, R.A. and Weed, D.L. 1998. Review of epidemiologic studies of alcohol and prostate cancer: 1971-1996. Nutrition and Cancer 30:1-13.

Breslow, R.A. and Ross, S.A., and Weed, D.L. 1998. Quality of reviews in epidemiology. Am J Public Health 88:475-477.

Weed, D.L. and Hursting, S.D. 1998. Biologic plausibility in causal inference: current method and practice. Am J Epidemiol 147:415-425.

Cronin, K.A., Weed, D.L., Prorok, P.C., and Connor, R.J. 1998. Case-control studies of screening: theory and practice. J National Cancer Inst. 90:498-504.

Weed, D.L. and McKeown, R.E. 1998. Epidemiology and virtue ethics. Int J Epidemiol. 27:343-348.

McKeown, R.E. and Weed, D.L. 1998. Authors' response to comment on "Epidemiology and virtue ethics." Int J Epidemiol. 27:348-349.

Huerta, E.E. and Weed, D.L. 1998. "Cuidando su Salud" (Spanish language radio in preventive medicine and public health). Cancer (Suppl) 83:1805-1808.

Weed, D.L. 1997. Methodologic guidelines for review papers. J Nat Cancer Inst. 89:6-7.

Weed, D.L. and Kramer, B.S. 1997. Response to Brind et al. (letter). J Nat Cancer Inst. 89:588.

Bulterys, M., Morgenstern, H., and Weed. D.L. 1997. Quantifying the expected versus potential impact of a risk-factor intervention program. Am J Public Health 87:867-868.

Weed, D.L. 1997. Underdetermination and incommensurability in contemporary epidemiology. Kennedy Institute of Ethics J. 7:107-127.

Weed, D.L. 1997. Meta-analysis under the microscope. J Nat Cancer Inst. 89:904-905.

Weed, D.L. 1997. The behavior-biology interface in cancer prevention and control. Prev Med. 26:S37-S41.

Merrill, R.M., Weed, D.L., and Feuer, E.J. 1997. The lifetime risk of developing prostate cancer in white and black men. Cancer Epidemiol Biom Prev, 6:763-768.

Weed, D.L. 1997. On the use of causal criteria. Int J Epidemiol. 26:1137-1141.

Gerlach, K.K., Marino, C., Weed, D.L., and Hoffman-Goetz, L. 1997. Lack of colon cancer coverage in seven women's magazines. Women & Health 26:57-68.

Weed, D.L. 1996. The sea of person time. Int J Epidemiol 25:1-4.

Weed, D.L. 1996. Book Review of: *The Fight for Public Health: Principles and Practice of Media Advocacy*. Prev Med, 25:86.

Weed, D.L. and Gorelic, L.S. 1996. The practice of causal inference in cancer epidemiology. Cancer Epidemiol Biomark Prev. 5:303-311.

Weed, D.L. 1996. Weed Responds to: The future of epidemiology: a humanist response. Am J Pub Health 86:1029-1030.

Weed, D.L. 1996. Book review of: Changing the Odds: Cancer Prevention through Personal Choice and Public Policy. Oncology 10:1432.

Weed, D.L. and Kramer, B.S. 1996. Induced abortion, bias and breast cancer: why epidemiology hasn't reached its limit. J Nat Cancer Inst. 88:1698-1700.

Weed, D.L. 1995. Epidemiology, the humanities, and public health. Am J Pub Health 85:914-918.

Houn, F., Bober, M.A., Huerta, E.A., Hursting, S., Lemon, S., and Weed, D.L. 1995. The association between alcohol and breast cancer: popular press coverage of research. Am J Pub Health 85:1082-1086.

Weed, D.L. 1994. Science, ethics guidelines, and advocacy in epidemiology. Ann Epidemiol. 4:166-171.

Reprinted by permission in: Coughlin SS (ed.). 1995. Ethics in Epidemiology and Clinical Research. Chestnut Hill, MA: ERI. Pp. 267-272.

Weed, D.L. 1994. Book Review of: Apricots and Oncogenes: On Vegetables and Cancer Prevention. Am J Epidemiol. 139:743-744.

Weed, D.L. 1994. Between science and technology: the case of antihistamines and cancer. J Nat Cancer Inst. 86: 740-741.

Coughlin, S.S., Benichou, J. and Weed, D.L. 1994. Attributable risk estimation in case control studies. Epidemiol Rev. 16:51-64.

Weed, D.L. 1994. Alcohol, breast cancer, and causal inference: where ethics meets epidemiology. Contemporary Drug Problems 21:185-204.

Greenwald, P.G., Kramer, B.K., and Weed, D.L. 1993. Expanding horizons in breast and prostate cancer prevention and early detection. J Cancer Education 8:91-107.

Husten, C., Weed, D.L, and Kaluzny, A.R. 1993. Training researchers in cancer prevention and control: a description and evaluation of NCI's Cancer Prevention Fellowship Program. J Cancer Education 8:281-290.

Pommerenke, F. and Weed, D.L. 1991. Physician compliance: Improving skills in preventive medicine practices. Am Fam Physician 43:560-568.

Weed, D.L. 1991. The merger of bioethics and epidemiology. J Clin Epidemiol. 44:15S-22S.

Cairoli, V.J. and Weed, D.L. 1991. NCI Report on the Cancer Education Program. J Cancer Education 6:65-66.

Connor, R.J., Prorok, P.C., and Weed, D.L. 1991. The case-control design and the assessment of the efficacy of cancer screening. J Clin Epidemiol. 44:1215-1221.

Clark, L.C., Patterson, B.H., Weed, D.L., and Turnbull, B.W. 1991. Design issues in cancer chemoprevention trials using micronutrients: application to skin cancer. Cancer Bulletin 43:519-524.

Koopman, J.S. and Weed, D.L. 1990. Epigenesis theory: A mathematical model relating causal concepts of pathogenesis in individuals to disease patterns in populations. Am J Epidemiol. 132:366-390.

Greenwald, P., Cullen, J.W., and Weed, D.L. 1990. Introduction: cancer prevention and control. Semin Oncol. 17:383-390.

Weed, D.L., Greenwald, P., and Cullen, J.W. 1990. The future of cancer prevention and control. Semin Oncol. 17:504-509.

Weed, D.L., Selmon, M., and Sinks, T. 1988. Links between categories of interaction. Am J Epidemiol. 127:17-27.

Weed, D.L. and Trock, B. 1988. Interactions and public health decisions. J Clin Epidemiol. 41:207-209.

Weed, D.L. 1987. The author replies. (Re: "On the logic of causal inference"). Am J Epidemiol. 126:157-158.

Weed, D.L. 1987. Epidemiology's triple crown. J Chronic Dis. 40:905-906.

Weed, D.L. 1987. The author replies. (Re: "On the logic of causal inference"). Am J Epidemiol. 126:557.

Weed, D.L., Tyroler, H.A., and Shy, C.M. 1987. The healthy worker effect in actively-working communications workers. J Occup Med. 29:335-339.

Weed, D.L. 1986. Historical roots of the healthy worker effect. J Occup Med. 28:343-347.

Weed, D.L. 1986. Lament for an epidemiologist. Pharos 49:43.

Weed, D.L. 1986. On the logic of causal inference. Am J Epidemiol. 123:965-979.

Weed, D.L. and Trock, B.J. 1986. Criticism and the growth of epidemiologic knowledge. (Re: "Popperian refutation in epidemiology"). Am J Epidemiol. 123:1119-1120.

Weed, D.L. 1985. An epidemiological application of Popper's method. J Epidemiol Commun. Health 39:277-285.

Weed, D.L. 1983. Ethics and chemoprevention research. Semin Oncol. 10:355-359.

BOOKS, BOOK CHAPTERS, EDITED JOURNAL ISSUES, DISSERTATION, AND TECHNICAL REPORTS

Weed, D.L. 2009. Towards a Philosophy of Epidemiology. In: Coughlin, S.S., Beauchamp, T.L., and Weed, D.L. (eds). Ethics and Epidemiology. New York:Oxford University Press.

Rockhill, B. and Weed, D.L. 2006. Increasing the contribution of epidemiology to the primary prevention of cancer. In Schottenfeld, D.A. and Fraumeni, J.F. Jr. (eds). Cancer Epidemiology and Prevention. Pp. 1292-1302.

Weed, D.L. 2004. Ethics and philosophy of public health. In Khushf. G. Handbook of Bioethics: A Philosophical Overview. Pp.525-547.

Weed, D.L. 2002. Philosophical basis for public health. In Breslow, L. et al. (ed). Encyclopedia of Public Health. Farmington Hill, Michigan: Macmillan Reference. Pp. 914-917.

Weed, D.L. 2002/2003. Is the Precautionary Principle a principle? IEEE Technology and Society Magazine 21:45-48.

[No authors listed]. American College of Epidemiology Ethics Guidelines. 2000. Ann Epidemiol 10(8):487-497.

WHO Working Group. 2000. Evaluation and use of epidemiological evidence for environmental health risk assessment: guideline document. World Health Organization, EUR/00/5020369, E68940. See also: Environmental Health Perspectives 108:997-1002, 2000.

Weed, D.L. 1999. Ethics and consent. In Kramer, B.S., Gohagan, J., and Prorok, P.C. (eds). Cancer Screening: Theory and Practice. New York: Marcel Dekker. Pp. 89-140.

Douglas Weed and B.S. Kramer, "Breast cancer studies aren't political," Wall Street Journal, Wednesday, 26 Mar 1997, sec. A19.

Weed, D.L. 1996. Epistemology and ethics in epidemiology. In Coughlin, S.S. and Beauchamp, T.L. Ethics and Epidemiology. New York: Oxford. Pp. 76-94.

Weed, D.L. 1995. Causal and preventive inference. In Greenwald, P., Kramer, B., and Weed, D.L. (eds.). Cancer Prevention and Control. New York: Marcel Dekker. Pp. 285-302.

Weed, D.L. and Coughlin, S.S. 1995. Ethics in cancer prevention and control. In Greenwald, P., Kramer, B., and Weed, D.L. (eds). Cancer Prevention and Control. New York: Marcel Dekker. Pp. 497-507.

Weed, D.L. and Husten, C. 1995. Training in cancer prevention and control. In Greenwald, P., Kramer, B., and Weed, D.L. (eds). Cancer Prevention and Control. New York: Marcel Dekker. Pp. 707-717.

Greenwald, P.G., Kramer, B.K., and Weed, D.L. 1995. Cancer Prevention and Control. New York: Marcel Dekker.

Greenwald, P.G., Cullen, J.W., and Weed, D.L. (eds.). Cancer prevention and Control. Seminars in Oncology 17:1990.

Weed, D.L. 1988. Causal criteria and Popperian refutation. In Rothman, K.J. (ed.). Causal Inference. Chestnut Hill: Epidemiology Resources, Inc. Pp. 15-32.

Weed, D.L. 1988. Criticism and its constraints: A self-appraisal and rejoinder. In Rothman, K.J. (ed.). Causal Inference. Chestnut Hill: Epidemiology Resources, Inc. Pp. 201-207.

Weed, D.L. 1982. An investigation of the healthy worker effect. Ph.D. dissertation, University of North Carolina, Chapel Hill, North Carolina.

ABSTRACTS

Weed, D.L. Case-based causality: an application of artificial intelligence to environmental carcinogenesis. AACR Special Conference on Environmental Carcinogenesis: Potential Pathway to Prevention. Charlotte, NC, June, 2019.

Weed, D.L. Cancer causation: rethinking the role of analogy. Society of Toxicology. San Antonio, Texas, March, 2018 and Ca Rsch 2017;77(Suppl):5301.

Weed, D.L., Alexander, D.D. A Systematic Review of the Use of the Bradford-Hill Criteria. Society of Toxicology. San Antonio, Texas, March, 2018.

Weed, D.L. Rethinking the role of analogy in causal inference. International Society of Environmental Epidemiology. Rome, Italy, September, 2016.

Kelsh, M.A., Yao, B., Arrindell, D.A., Alexander, D., Acquavella, J., Weed, DL. Evaluation of adverse events of pharmaceutical agents—the impact of effect measure selection on meta-analysis findings. International Society of Pharmacoepidemiology. October, 2014. Taipei, Taiwan.

Weed, D.L., Althuis, M.A. Evaluating Confounding Bias when Designing Meta-analyses of Dietary Risk Factors with Weak Associations: a Systematic Review of Risk Factors for Type 2 Diabetes. International Epidemiology Association World Congress. August, 2014. Anchorage, Alaska.

Weed, D.L. 2012. On the utility of criteria-based methods of causal inference. International Society for Environmental Epidemiology: Conference. August, 2012.

Weed, D.L. 2009. Meta-analysis and causal inference: a case study of benzene and non-Hodgkin's lymphoma. Benzene 2009: Health Effects and Mechanisms of Bone Marrow Toxicity, Implications for t-AML and the Mode of Action Framework. Munich, Germany.

Weed, D.L. 2008. A method for individual causation. Am. J. Epidemiol. 167:S115.

Weed, D.L. 2001. Ethics of precautionary preventive interventions. Am. J. Epidemiol. 153:S7.

Weed, D.L. 2001. Environmental epidemiology: basics and proof of cause-effect. Toxicology 164:29.

Weed, D.L. 2000. Epidemiology, beneficence, and the Precautionary Principle. Am J. Epidemiol. 151: S90.

Marcus, P.M. and Weed, D.L. 1999. Did the chronic disease era of cancer epidemiology start before we think it did? Ann. Epidemiol.

Weed, D.L. 1997. Underdetermination and incommensurability in epidemiology. Am. J. Epidemiol. 145(Suppl):S73.

Weed, D.L. 1993. Bioethical methods in epidemiology. Am. J. Epidemiol. 138:671.

Weed, D.L. 1992. Epidemiology and the humanities: an illustrated example. Am. J. Epidemiol. 136:1006-1007.

Gorelic, L.S., Weed, D.L., Dresser, C., Graubard, B., and Ruiz, E. 1991. Cervical cancer screening practices by Hispanic women. Prev. Med.

Weed, D.L. 1991. Causal inference: A matter of principle. Am. J. Epidemiol. 134:779-780.

Patterson, B.H., Clark, L.C., and Weed, D.L. 1991. Cancer prevention trials using micronutrients: design issues. Cont. Clin. Trials.

Weed, D.L. and Gorelic, L. 1989. Weak associations, bias, and causal inference. Am. J. Epidemiol. 130:819.

Mayer, W.J., Weed, D.L., and Trock, B.J. 1989. Criteria for preventive inference. Am. J. Epidemiol. 130:854.

Chu, K.C. and Weed, D.L. 1989. Validating screening case-control study results with randomized controlled trials for screening. Am. J. Epidemiol. 130:826-827.

Weed, D.L. 1988. Analyzing conflicts of interest in epidemiologic research. Am. J. Epidemiol. 128:943.

Weed, D., Connor, R., and Prorok, P. 1986. Case-control studies of screening: A methodological test. Am. J. Epidemiol. 124:527.

Weed, D. and Trock, B. 1985. Preventive interactions. Am. J. Epidemiol. 122:509-510.

Weed, D. 1985. Causal criteria and Popperian refutation. Am. J. Epidemiol. 122:550.

Trock, B and Weed, D. 1985. Predicting the effects of retinoid chemoprevention. Am. J. Epidemiol. 122:521-522.

Weed, D.L., Selmon, M., and Sinks, T. 1984. Predicting interactions. Am. J. Epidemiol. 120: 464-465.

Weed, D.L. 1983. An epidemiologic application of Popper's method. Am. J. Epidemiol. 118:432.

Weed, D.L. 1982. Age and the healthy worker effect: New findings with old measures. Am. J. Epidemiol. 116:574-575.

PRESENTATIONS:

A Mortality Study in Communications Workers. Society for Epidemiologic Research Student Workshop on Methods, Minneapolis, Minnesota, June, 1980.

Age and the healthy worker effect: new findings with old measures. 15th Meeting of the Society for Epidemiologic Research, Cincinnati, Ohio, June, 1982.

Absolute and relative measures of effect. National Cancer Institute, Division of Cancer Prevention and Control, Bethesda, Maryland, November, 1982.

An epidemiologic application of Popper's method. National Cancer Institute, Division of Cancer Prevention and Control, Bethesda, Maryland, May, 1983.

An epidemiologic application of Popper's method. 16th Meeting of the Society for Epidemiologic Research, Winnipeg, Manitoba, Canada, June, 1983.

Ethics and chemoprevention. National Cancer Institute, Division of Cancer Prevention and Control, Bethesda, Maryland, July, 1983.

Epidemiology and the engineer. 36th Annual Conference on Engineering in Medicine and Biology, Columbus, Ohio, September, 1983.

Disease models and inference in epidemiology. National Meeting of the Operations Research Society of America, Orlando, Florida, November, 1983.

Disease models and epidemiologic inference. Department of Preventive Medicine, Cornell University, Ithaca, New York, November, 1983.

Popper and epidemiology. Division of Preventive Medicine, Walter Reed Army Institute of Research, Washington, D.C., April, 1984.

Some issues in predicting interactions. National Cancer Institute, Division of Cancer Prevention and Control, Bethesda, Maryland, May, 1984.

Predicting interactions. 17th Meeting of the Society for Epidemiologic Research, Houston, Texas, June, 1984.

Cancer control epidemiology. Ohio State Comprehensive Cancer Center, Columbus, Ohio, August, 1984.

Modelling disease interactions. 37th Annual Conference on Engineering in Medicine and Biology, Los Angeles, California, September, 1984.

Causal Criteria and Popperian Refutation. 18th Annual Meeting of the Society for Epidemiologic Research, Chapel Hill, North Carolina, June, 1985.

Preventive Interactions. 18th Annual Meeting of the Society for Epidemiologic Research, Chapel Hill, North Carolina, June, 1985.

Case-control studies of screening: A methodologic test. 19th Annual Meeting of the Society for Epidemiologic Research, Pittsburgh, Pennsylvania, June, 1986.

Speaking in Tongues: A Mega-analysis of a debate. National Cancer Institute, Division of Cancer Prevention and Control, Bethesda, Maryland, July, 1987.

The Analysis of Debates and Other Forms of Epidemiologic Reasoning. University of Pennsylvania School of Medicine, Clinical Epidemiology Unit, Philadelphia, Pennsylvania, January, 1988.

The Analysis of Medical Reasoning. Southern Illinois University School of Medicine, Springfield, Illinois, February, 1988.

Interaction. Uniformed Services University of the Health Sciences, Division of Preventive Medicine and Biometrics, Bethesda, Maryland, May, 1988.

Modelling Interactions in Epidemiologic Research. University of Michigan, School of Public Health, Department of Epidemiology, Ann Arbor, Michigan, May, 1988.

Analyzing Conflicts of Interest in Epidemiologic Research. 21st Annual Meeting of the Society for Epidemiologic Research, Vancouver, British Columbia, Canada, June, 1988.

Analyzing Conflicts of Interest. Office of Protection from Research Risks, Office of the Director, NIH, Bethesda, Maryland, July, 1988.

Epidemiology and the Ethics of Prevention. The Johns Hopkins University, School of Hygiene and Public Health, Department of Epidemiology, Baltimore, Maryland, November, 1988.

Ethical Problems in Cancer Prevention. The Johns Hopkins University, School of Hygiene and Public Health, Department of Epidemiology, Baltimore, Maryland, January, 1989 and June, 1989.

The Future of Cancer Prevention and Control. The University of North Carolina, Lineberger Cancer Center, Chapel Hill, North Carolina, March, 1989.

On the Merger of Bioethics and Epidemiology. IEF Conference on Ethics in Epidemiology, Birmingham, Alabama, June, 1989.

Weak Associations, Bias, and Causal Inference. 22nd Annual Meeting of the Society for Epidemiologic (SER), Birmingham, Alabama, June, 1989.

Criteria for Preventive Inference. Centers for Disease Control, Atlanta, Georgia, September, 1989.

Causal Inference. University of Virginia, College of Medicine, Charlottesville, Virginia, October, 1989.

Uniformed Services University of the Health Sciences, Bethesda, Maryland, May, 1991, May, 1992, April 1993, April 1994.

Ethics in Epidemiology. University of Maryland at Baltimore, College of Medicine, Baltimore, Maryland, December, 1989.

Science, Ethics and the Prevention of Cancer. Fox Chase Cancer Center, Philadelphia, Pennsylvania, December 1989.

Ethics and Cancer Prevention. National Cancer Institute, Bethesda, Maryland, December, 1989.

Common Sense in Epidemiology. The Johns Hopkins University, School of Hygiene and Public Health, Department of Epidemiology, Baltimore, Maryland, January, 1990.

Centers for Disease Control, National Institute of Occupational Safety and Health, Robert A. Taft Laboratories, Cincinnati, Ohio, February, 1990.

Yale University School of Medicine, Department of Epidemiology and Public Health, New Haven, Connecticut, April, 1990.

University of Virginia, College of Medicine, Division of Epidemiology and Virology, Charlottesville, Virginia, October, 1990.

University of North Carolina, School of Public Health, Department of Epidemiology, Chapel Hill, North Carolina, November, 1991.

Harvard University, School of Public Health, Department of Epidemiology, Boston, Massachusetts, December, 1991.

Case Studies in Epidemiological Ethics. Yale University School of Medicine, Department of Epidemiology and Public Health, New Haven, Connecticut, April, 1990.

Inferential Issues in the Study of Alcohol and Breast Cancer. AMC Cancer Research Center, Denver, Colorado, December, 1990.

Epidemiology and the Humanities. Society for Health and Human Values, St. Louis, Missouri, October, 1991.

Bringing Ethics into Causal Inference in Epidemiology. University of Virginia, College of Medicine, Division of Epidemiology and Virology, Charlottesville, Virginia, April, 1992.

Training Programs for Cancer Prevention and Control Researchers. NCI Cancer Center Directors' Workshop, Buffalo, New York, June 1992.

Science, Ethics, and Public Policy: The Case of Alcoholic Beverages and Breast Cancer. American College of Epidemiology, Bethesda, Maryland, September 1992.

Cancer Prevention Training in the Preventive Oncology Branch. American Cancer Society Board of Directors Meeting, Atlanta, Georgia, November 1992.

Ethical Issues in Prophylactic Mastectomy. American Society of Preventive Oncology, Tucson, Arizona, March 1993.

Ethical Considerations in Moderate Alcohol Drinking. Addiction Research Foundation International Symposium, Toronto, Ontario, May 1993.

Ethics in Epidemiology. PHS Epidemiology Training Program Seminar, Bethesda, Maryland, August 1993, and Uniformed Services University of the Health Sciences, Basic Epidemiology I, Bethesda, Maryland, November 1993, November 1994, November 1995.

Alcohol and Breast Cancer. University of Hawaii Cancer Center, Honolulu, Hawaii, September 1993.

Public Health Posters from the National Library of Medicine. Society for Health and Human Values, Rosslyn, Virginia, November 1993.

Untangling Decisions in Health Matters: The Case of Cancer Prevention. Baltimore Ethical Society, Baltimore, Maryland, April 1994.

Evidence-based Cancer Prevention: How Do We Know What to Do? Ohio State University Preventive Medicine Alumni Conference, Columbus, Ohio, September 1994.

Preventive Medicine and the Press: A Case Study. Ohio State University Preventive Medicine Alumni Conference, Columbus, Ohio, September 1994.

Causal Inference in Cancer Epidemiology: A Methodologic Review. Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts, September 1994.

Division of Cancer Prevention and Control, NCI, Bethesda, Maryland, September 1994.

Department of Epidemiology, University of Washington School of Public Health and Community Medicine, Seattle, Washington, September 1995.

Epidemiology, Humanities, and Public Health. Department of Epidemiology and Preventive Medicine, University of Maryland, Baltimore, Maryland, September 1994.

Should Epidemiologists be Advocates? Centers for Disease Control and Prevention Course on Government Employees and Public Policy, Atlanta, Georgia, January 1995.

A New Ethic for Epidemiology? Third Brazilian Congress of Epidemiology, Salvador, Bahia, Brazil, April 1995.

Causality, Data and Inference. George Washington University School of Medicine, Washington D.C., May 1995.

The Future of Epidemiology. Uniformed Services University of the Health Sciences, Bethesda, MD, August 1995, August 1996, August 1997.

Beyond Black Box Epidemiology: Behavior and Biology. American College of Epidemiology Annual Meeting, St. Louis, MO, September 1995.

Causal Conclusions, Public Health Recommendations and Methods of Ethical Reasoning: A Practical Approach. American Public Health Association Annual Meeting, San Diego, CA, October 1995.

Biologic Evidence and Human Cancer Causation. Department of Epidemiology, MD Anderson Cancer Center, Houston, TX, March 1996.

Epidemiology Branch, National Institute for Environmental Health Sciences, Research Triangle Park, NC, July 1996.

American Health Foundation, Valhalla, NY, January 1998. Department of Oncology, McGill University, Montreal, QC, Canada, February 1999.

Preventing Scientific Misconduct. Department of Epidemiology and Biostatistics, Tulane University School of Public Health and Tropical Medicine, New Orleans, LA, March 1996.

Preventive Medicine Residency Program, Centers for Disease Control and Prevention, Atlanta, GA, March 1996.

Department of Chemistry, University of Maryland, College Park, MD, April 1996.

University of Hawai'i at Manoa, Honolulu, HI, August 1996.

Department of Biometrics and Preventive Medicine, Uniformed Services University of the Health Sciences, Bethesda, MD, November 1996 and November 1997.

MD Anderson Cancer Center, Houston, TX, July 1998.

Department of Oncology, Royal Victoria Hospital, McGill University, Montreal, QC, Canada, February 1999.

Office of Research Integrity, U.S. Public Health Service, Rockville, Maryland, September 2000.

Epidemiology and Virtue Ethics. XIVth Congress of the International Epidemiological Association, Nagoya, Japan, August 1996.

Association or Causation: Myths and Legends. NIH Research Festival Workshop, Bethesda, MD, September 1996.

On the Need for Ethics in the Community of University Scholars. University of South Carolina, Columbia, SC, March 1997.

Communicating Cancer Information: an American Perspective. German Cancer Research Center, Heidelberg, Germany, April 1997.

Annual Meeting of the European Association for Cancer Education, Brussels, Belgium, April 1997.

Women's Health and the Media: The Role of the Medical Journal. Healthy Women 2000 Conference, Washington DC, June 1997.

Principles and Practice of Cancer Prevention and Control. NCI Medical Oncology Lecture Series, Bethesda, MD, August 1997.

Ethics and Cancer Screening. Cancer Conference: Integrating Public Health Programs for Cancer Control, Atlanta, GA, September 1997.

NIH Research Festival. Bethesda, MD, October 1997.

University of Puerto Rico, San Juan, Puerto Rico, February 1998.

Pavilion du Chum, University of Montreal, Montreal, QC, Canada, February 1999.

Publishing and Authorship. Cancer Prevention Fellows Data Club Meeting, Bethesda, MD, September 1997.

Towards a Philosophy of Epidemiology. Department of Social and Preventive Medicine, SUNY-Buffalo, Buffalo, NY, September 1997.

Causal Criteria in Nutritional Epidemiology. ILSI Conference on the Role of Epidemiology in making Nutritional Recommendations. Washington, D.C., October 1997.

Philosophical Foundations for the Practice of Epidemiology. III Congress of the Chilean Society of Epidemiology. Vina del Mar, Chile, October 1997.

16th Meeting of the Spanish Society of Epidemiology. Sevilla, Spain, October 1998.

Determining Causality from Epidemiological Studies.

III Congress of the Chilean Society of Epidemiology. Vina del Mar, Chile, October 1997.

Department of Epidemiology and Biostatistics, Boston University School of Public Health, Boston, MA, December 1997.

University of Puerto Rico, San Juan, Puerto Rico February 1998.

Department of Epidemiology and Biostatistics, McGill University, Montreal, QC, Canada, February 1999.

End of the Era of Weak Associations? An Historical Study of Epidemiologic Discovery. NIH Historical Office Symposium on Evidence and Action: How epidemiologists make decisions about science and the public's health. NIH Clinical Center, Bethesda, MD, March 1998.

Department of Epidemiology. School of Public Health. University of North Carolina, Chapel Hill, NC, April 1998.

Center for Clinical Epidemiology and Biostatistics. University of Pennsylvania Medical Center, Philadelphia, PA, May 1998.

MD Anderson Cancer Center. Houston, TX, July 1998.

Department of Epidemiology and Biostatistics, Yale University, New Haven, CT, November 1998.

Department of Epidemiology, University of California, Berkeley, Berkeley, CA, March 1999.

Channing Lab, Harvard University, May 1999.
Department of Health Evaluation Sciences, University of Virginia, Charlottesville, VA, January 2000.

Department of Epidemiology, Michigan State University, East Lansing, MI, February 2000.

Department of Epidemiology, Tulane University School of Public Health and Tropical Medicine, New Orleans, LA, April 2000.

Epidemiology at a Crossroads. Seminar for the Cosmetics, Fragrance, and Toiletries Association. Morristown, NJ, April 1998.

Epidemiologists and Risk: Theory, Method, and Practice. Workshop on Epidemiology and Toxicology. Washington, DC, May 1998.

The Interpretation of Meta-Analyses with reference to Causal Inference and Public Health Decisionmaking. Society for Epidemiologic Research Symposium on the Methods and Applications of Meta-Analysis. Chicago, IL, June 1998.

Roles and Responsibilities of Epidemiologists. Department of Epidemiology, University of North Carolina, Chapel Hill, NC, March 1999.

Causation and Biology. Society for Epidemiologic Research Symposium on the Future of Causes in Epidemiology. Baltimore, MD, June 1999.

Epidemiologic Evidence and the Precautionary Principle. International Society for Environmental Epidemiology. Athens, Greece, September 1999.

Improving Cancer Screening: An American Perspective. Symposium on Cancer Screening. Catholic University of Korea Cancer Center. Seoul, Korea, October 1999.

Causality and Inference in Cancer Epidemiology: We've Got Some Problems.

Ohio State University James Cancer Hospital. Columbus, OH, October 1999.

Department of Food Science and Human Nutrition. Michigan State University, East Lansing, MI, February 2000.

Department of Epidemiology. Tulane University School of Public Health and Tropical Medicine, New Orleans, LA, April 2000.

Our future is not epidemiology.calm. American Public Health Association Special 70th Anniversary for the Epidemiology Section. Chicago, IL, November 1999.

American College of Epidemiology Ethics Guidelines: Foundations and Dissemination. AAAS Conference on Research Integrity, Washington, DC, April 2000.

Teaching Ethics and Public Health: Curriculum Content. ASPH/HRSA Workshop on Ethics and Public Health, Washington, DC, May 2000.

Science, Ethics and the Future of Preventive Oncology. Seminars in Clinical and Molecular Oncology, National Cancer Institute, Bethesda, MD, July 2000.

Precautionary Principle and the Philosophy of Public Health. WHO Workshop on the Precautionary Principle, Rome, Italy, May 2001.

Science, Ethics and the Future of Epidemiology.

International Epidemiological Association Regional Asia Meeting, Kitakyushu, Japan, September 2001.

Kyoto University School of Public Health, Kyoto, Japan, September 2001.

National Cancer Center, Tokyo, Japan, September 2001.

Cancer Prevention in the 21st Century Istituto Superiore di Sanita, Rome, Italy, May 2002.

Roles and Responsibilities of Epidemiologists Istituto Superiore di Sanita, Rome, Italy, May 2002.

Promoting Research Integrity Cleveland Clinic Foundation, Cleveland, Ohio, May 2002.

Scope and Importance of Public Health World Bank/WHO Conference on Public Health Challenges in the Middle East and North Africa, Beirut, Lebanon, June 2002.

The Precautionary Principle and the Philosophy of Public Health International Society of Environmental Epidemiology, Vancouver, BC, August 2002.

Appendix F

An Assessment of the Proposed Relationship Between Lead Exposure and Kidney Function in Children

Introduction

It has been alleged that the plaintiffs in the "Minors Subclass" are likely to suffer renal disease resulting from their alleged lead exposure during the Flint Water Switch. I have been asked to assess the proposed relationship between lead exposure and kidney function (including kidney disease) in children and as adults (given childhood lead exposure of the type potentially experienced during the Flint Water Switch). As discussed earlier in this report, capillary tests tend to overestimate blood lead levels. Hence, the issue here is whether a general causation relationship exists between very low levels of lead—at levels similar to those estimated in the plaintiffs or less—and renal function, including renal diseases.

The approach taken here is to systematically identify, describe, and interpret relevant epidemiological studies and reviews that have examined or commented on the relationship between blood lead levels and renal function. The ATSDR Report on the Health Effects of Lead (2020) provided one source for identifying relevant studies and reviews; that report identified the following studies: Staessen et al. (1992), de Burbure et al. (2006), Fadrowski et al. (2010), and Khan et al. (2010). An additional search of PubMed on November 19, 2020 identified Fadrowski et al. (2013); details of that search are available upon request. Careful examination of reference lists of these studies and searches of other databases (e.g. Google Scholar) identified the following relevant publications: Bernard et al. (1995); Fels et al. (1998), Katnelson et al. (2007), Loghman-Adham et al. (1997), Rubio-Andrade et al. (2016), Sanders et al. (2019), Skroder et al. (2016), Weidemann et al. (2016), and Zheng et al. (2017).

A brief description of these publications follows.

Reviews of the Renal (Kidney) Effects of Lead Exposure in Children

Loghman-Adham (1997) is an early review that discusses the relationship between environmental lead exposure and renal function (including renal disease). The studies identified and discussed in this review involve extraordinarily high blood lead levels relative to those measured in the plaintiffs from Flint, Michigan. These studies included Tepper (1963), Chisolm (1970), Moel and Sachs (1992), Hu (1991), Payton et al. (1994), Verberk et al. (1996), and McDonald et al. (1996). Logham-Adham (1997) discussed some studies that "suggested" a relationship between childhood lead exposure and subsequent impairment of renal function. Nevertheless, many of the studies reviewed by Loghman-Adham (1997) revealed that children who had been exposed to high levels of lead and who had very high blood lead levels had no evidence of lead nephropathy—i.e. kidney disease—later in life. The study by (Moel and Sachs, (1992), for example, revealed children with blood lead levels on the order of > 100 µg/dL and thus almost 100 times higher than any Flint, Michigan plaintiff. The study by Chisolm (1970) had similar results. As Loghman-Adham (1997, p. 931) writes, Chisolm "found no evidence of nephropathy" in children with known lead poisoning, 11-16 years after the event. The study by Hu (1991) "found no evidence of impaired renal function, as assessed by measurement of glomerular filtration" (Logham-Adham, 1997, p. 931). The study by Payton et al. (1994) as discussed by Logham-Adham (1997) observed similar findings; in this study, 454 individuals with a history of childhood lead poisoning (i.e. with blood lead levels between >80 and 100 µg/dL) between 1923 and 1966 had no evidence of chronic nephritis as a cause of death.

Weidemann et al. (2016) is a broad review discussing studies of several environmental exposures and kidney health in children and adolescents. The metals and chemicals discussed include but are not limited to arsenic, cadmium, lead, mercury, and uranium. The authors describe the appearance of chronic lead nephropathy in young adults whose exposure to lead revealed blood lead levels > 60 µg/dL, i.e. at levels much higher than any Flint, Michigan plaintiff. The authors note that early studies—similar to those reviewed by Logham-Adham (1997) above—"have not demonstrated overall increased mortality of statistically significant decreased kidney function" (Weidemann et al. 2016, p. 6). More recent studies of the potential effects of lead on the kidney have involved populations living near industrial pollution sources and the authors note that the results have been "contradictory" (Weidemann et al. 2016, p. 6). Finally, the authors note that reverse causality is a reasonable explanation for results of studies that show a relationship between higher lead levels and reduced kidney function, especially in cross-sectional studies. Why? Because in children with reduced kidney function for other reasons, there will be a reduction in the excretion of lead and therefore higher blood lead levels.

Zheng et al. (2017), like Weidemann et al. (2016) examines several environmental exposures and pediatric kidney function and disease. A prominent difference is that the Zheng et al. (2017) is a systematic review. The authors identified 16 studies that examined the possible relationship between exposure to lead and kidney outcomes; the studies were published between 1985 and 2016. There were 12 cross-sectional studies, 2 prospective cohorts, and 2 case-control studies. Focusing primarily upon the 2 prospective studies, the authors note that these studies "had larger study populations and adjusted for possible confounders" and were of high quality (Zheng et al. 2017, p. 640). These studies failed to find an association between blood lead levels and chronic kidney disease or a reduction in glomerular filtration rate (Fadrowski et al. 2010, 2013). It follows that there is no convincing evidence that low blood lead levels in children causes chronic kidney disease in later life.

In sum, the published peer-reviewed reviews on relatively low blood lead levels and kidney function (or kidney disease) found no association when the studies were prospective and of high quality. These are the studies that need to be considered in light of the known blood lead levels of the Flint, Michigan plaintiffs.

Organizational Reports

The **National Toxicology Program's Monograph on the Health Effects of Low-Level Lead** (2012, p. xxii) concludes that there is "inadequate evidence to address the potential association between blood Pb levels < 10 µg/dL in children < 12 years of age and impaired kidney function, because results are inconsistent and available studies of kidney function in young children are less reliable in general because tests of kidney function lack clear predictive value in this age group." The same report notes that there is a single cross-sectional study of blood Pb levels < 5 µg/dL and adverse effects on kidney function in children < 12 years of age. According to the NTP, a single study—lacking prospective data—is considered "limited." In any event, the NTP does not conclude that low levels of lead cause decrements in kidney function, much less kidney disease in children.

The **Agency for Toxic Substances and Disease Registry (ATSDR) Toxicological Profile for Lead (2020)** describes four cross-sectional studies of children exposed to lead and kidney parameters (e.g. serum creatinine, glomerular filtration rate, and serum cystatin). One study reveals that blood lead levels between 3.64 and 6.51 µg/dL are associated with an improvement in kidney function as measured by

serum creatinine (de Burbure et al. 2006). Three additional cross-sectional studies show modest decreases in kidney function with increasing blood lead levels (Khan et al. 2010; Fadrowski et al. 2010; and Staessen et al. 1992). The ATSDR report also makes clear that these studies, in addition to fact they are not prospective, fail to control for important confounding factors, including age, blood pressure, and exposure to other nephrotoxicants (e.g. cadmium). In the end, the ATSDR (2020) provides limited evidence for a relationship between childhood lead exposure and reduced kidney function and no (zero) evidence that childhood lead exposure is associated with chronic kidney disease later in life.

Sanders et al. (2019) is a recent cross-sectional study not included in the reviews and organizational reports described above has appeared in the peer-reviewed literature. It is important to point out that this study does not carefully separate the potential effects of lead, cadmium, mercury, and arsenic, but rather analyzes these metals as mixtures. For example, the authors defined a parameter "U_{mix}" as the urine levels of arsenic (As), cadmium (Cd), lead (Pb) and mercury (Hg). They also defined a parameter "B_{mix}" as the combination of Cd, Pb, and Hg in blood. The authors used the NHANES dataset for their cross-sectional analyses, subject to reverse causality. They reported that the U_{mix} parameter was associated with higher BUN, higher eGFR, and higher urine albumin but that these increases were primarily driven by mercury (Hg) and cadmium (Cd). For the B_{mix} parameter, the authors found a "marginally significant" relationship between the parameter and serum uric acid but no association between B_{mix} and urine albumin, eGFR, BUN, or blood pressure. In the end, this study revealed no clear evidence that lead is associated with any kidney health parameter.

Summary

It would be scientifically inappropriate to conclude from the evidence reviewed above that exposure to lead in childhood causes renal disease in adulthood. Similarly, it would be scientifically inappropriate to conclude that exposure to lead in childhood at levels similar to those measured in the Flint plaintiffs causes reductions in kidney function. The plaintiffs' expert's claim that the Flint plaintiffs are likely to suffer kidney disease as adults is pure speculation, devoid of scientific evidence.

To put these conclusions on an objective basis, I will apply the Bradford Hill criteria—described earlier in this report—to the available evidence on low levels of lead and renal disease (including renal function). An effort to apply the Bradford-Hill criteria to this evidence will fail. Nearly all studies are cross-sectional and cannot therefore contribute to a causal analysis. Furthermore, the results are inconsistent, the measured "effects" are weak, and there is no clear evidence of a dose response relationship. Indeed, it is not clear that an association exists "beyond that which we would attribute to chance" between lead and increased blood pressure in children. In addition, there is no valid observed association between lead and renal dysfunction or renal disease in children. In the absence of an association, applying the Bradford-Hill criteria is scientifically inappropriate.

References

Agency for Toxic Substances and Disease Registry (ATSDR) Toxicological Profile for Lead. 2020.

Logham-Adham M. Renal effects of environmental and occupational lead exposure. Environ Health Perspect 1997;105:928-38.

National Toxicology Program (NTP). Monograph on Health Effects of Low-Level Lead. U.S. Department of Health and Human Services, 2012.

Sanders AP, Mazzella MJ, Malin AJ, et al. Combined exposure to lead, cadmium, mercury, and arsenic and kidney health in adolescents age 12-19 in NHANES 2009-2014. Environ Int 2019;131:10.1016/j.envint.2019.104993.

Weidemann DK, Weaver VM, Fadrowski JJ. Toxic environmental exposures and kidney health in children. Pediatr Nephrol 2017;31:2043-54.

Zheng LY, Sanders AP, Saland JM, et al. Environmental exposures and pediatric kidney function and disease: A systematic review. Environ Res 2017;158:625-48.

Appendix G

An Assessment of the Proposed Relationship Between Lead Exposure and Hypertension (i.e. Increased Blood Pressure) in Children

A Systematic Review of Lead Exposure, Blood Pressure, and Hypertension in Children

Literature Search and Included/Excluded Published Papers

A systematic search of the medical literature was performed for publications examining the hypothetical relationship between exposure to lead and increased blood pressure as well as hypertension in children. The basic purpose of this search was to identify all English-language epidemiological studies and meta-analyses of epidemiological studies that have examined these hypothetical relationships. The specific aim of the search was to include all relevant studies and meta-analyses published to the present with emphasis on publications since the Agency for Toxic Substances and Disease Registry (ATSDR) final report on lead (2020). In addition, relevant reviews, reports, and commentaries were identified including some published prior to 2018 (the last year ATSDR included studies regarding lead and blood pressure in children in its 2020 report). Excluded were: letters to the editor, case reports (or case series) and textbook chapters addressing aspects of these hypothetical relationships. Reports of other organizations—and websites of same—representing information from the so-called "grey" literature were also examined, including the US EPA's (2014) review and the ATSDR's (2019) review.

Medline (PubMed) and EMBASE were searched for English-language publications as follows:

PubMed	September 16, 2020: search terms: "lead" and "blood pressure" and "children" Limits: 2018 to present. (n = 31)
EMBASE	September 18, 2020: search terms: "lead" and "blood pressure" and "children" (n = 298)

After removal of duplicates (n =12), a total of three hundred seventeen (317) potentially relevant publications were available for review.

After examining the titles and abstracts, a total of six (6) publications were identified for full text review, all published in or after 2018 and thus were potentially relevant: Ahn et al. (2018), Farzan et al. (2018), Sanders et al. (2018), Zachariah et al. (2018), Siddiqui and Maltesta (2020), and Yao et al. (2020).

The three hundred and ten (310) articles excluded were on topics other than the relationship between lead exposure and blood pressure in children. However, identified in these searches were several publications describing the general phenomenon of increased blood pressure and hypertension in children as well as other chemicals known or believed to be risk factors for increased blood pressure in children. These additional publications were as follows: Falkner (2010), Flynn et al. (2017), Flynn and Falkner (2017), Chen et al. (2019), Dong et al. (2019), Madhloum et al. (2019), Sander et al. (2019), and Sol et al. (2020).

An additional seven articles were identified for full text review from reference lists of these publications that did not appear in the ATSDR or US EPA Reports: Ahn et al. (2018), Camaj et al. (2018), Farzan et al.

(2018), Sanders et al. (2018), Zachariah et al. (2018), Dong et al. (2019), Siddiqui and Malatesta (2020) and Yao et al. (2020).

Additional relevant publications may emerge if more recent searches are attempted.

Studies and Reviews of the Hypothetical Relationship Between Lead Exposure and Blood Pressure (and Hypertension) in Children Published in or after 2018

Background: Normal and Elevated Blood Pressure Plus Hypertension (HTN) in Children and Adolescents

High blood pressure in children and adolescents has been increasing in prevalence and incidence in the United States due in large part to the rise in childhood obesity (Falkner, 2010). Another primary cause of childhood hypertension is heredity, i.e. family history (Flynn et al. 2017). So-called secondary causes of childhood hypertension include renal disease, renovascular disease, cardiac abnormalities (e.g. coarctation of the aorta), endocrine disorders, prematurity and low birthweight, and some pharmaceutical agents (e.g. oral contraceptives, central nervous system stimulants, and corticosteroids). Environmental agents (e.g. lead, cadmium, mercury, and phthalates) have been listed by some authors as "associated with childhood hypertension" (Siddiqui and Malatesta, 2020) but for lead this conclusion is based on a single cross-sectional study by Gump et al. (2007) as described by Flynn et al. (2017) ignoring the much larger body of evidence that will be discussed in more detail below.

It is important to distinguish between elevations in blood pressure and frank hypertension. The American Academy of Pediatrics has defined normal blood pressure, elevated blood pressure (previously called "prehypertension") and hypertension as the following:

> Normal Blood Pressure in Children and Adolescents
> Pre-adolescents: blood pressure ≤ 90th percentile of age/height/gender matched "controls" of normal weight
> Adolescents: (≥ 13 years): blood pressure ≤ 120/80
>
> Prehypertension in Children and Adolescents
> Pre-adolescents: blood pressure between 90th and 95th percentile of age/height/gender matched "controls" of normal weight
> Adolescents: blood pressure > 120/80 to 129/80
>
> Hypertension in Children and Adolescents (Stage I)
> Pre-adolescents: blood pressure ≥ 95th percentile of age/height/gender matched "controls" of normal weight
> Adolescents: blood pressure: 130/80 to 139/89

Important Note Regarding the Scope of this Systematic Review

As noted above, the purpose of this review is to systematically examine the studies of lead exposure and blood pressure in children published in 2018 (or later) that did not appear in the ATSDR report (2020). The ATSDR report (2020, p. 3) concluded the following after reviewing several earlier studies (Gerr et al. 2002), Gump et al. (2005), Gump et al. (2011), Zhang et al. (2011) and Ahn et al. (2018):

"A few studies show increased blood pressure in children and pregnant women."

Later in the report however, the ATSDR (2020, p. 63) echoes the fact that there are few studies and that the studies tend to be small when compared to studies in adults.

Importantly, the studies reviewed by ATSDR show inconsistent results. ATSDR only notes that "two prospective studies **suggest** that prenatal exposure to Pb is associated with increased blood pressure in childhood (Gump et al. 2005, Zhang et al. 2011)" *emphasis added.* On the other hand, the ATSDR also notes that "no association between PbB and diastolic or systolic blood pressure or risk of prehypertension (was observed) in a larger population of adolescents (n=1,776) with a mean PbB of 1.19 µg/dL (Ahn et al. 2018)." Importantly, the ATSDR did not make any claim that exposure to lead was associated with much less caused frank hypertension in children.

ATSDR makes no causal claim. In fact, ATSDR emphasizes the inconsistency and uncertainty of any relationship involving lead exposure and hypertension in children.

Given the clear inconsistency and uncertainty of any relationship between lead exposure and increased blood pressure in children and adolescents in these earlier studies and reports, it will prove important to examine the results of studies published in and after 2018. As noted above, these include Farzan et al. (2018), Sanders et al. (2018), Zachariah et al. (2018), and Yao et al. (2020). Dong et al. (2019) and Siddiqui and Malatesta (2020) are not studies per se, but rather general discussions of the epidemiology and clinical characteristics of blood pressure elevations in children.

Farzan et al. (2018) is the report of a prospective study of prenatal lead and arsenic exposure and blood pressure in children. The study—the New Hampshire Birth Cohort (NHBCS—was designed to examine the relationship between maternal Pb and As levels and blood pressure for women who reported using water from a water well at their residence. The authors examined maternal prenatal (28 weeks gestation) and 6-weeks postpartum toenail lead and arsenic levels. Blood pressure levels in children 5.5 years later were used in the analyses. Data on medical and lifestyle characteristics were collected by questionnaire; included were sociodemographic factors (age, race/ethnicity, marital status, education), reproductive history, and health history including tobacco and alcohol use. At 5 years, the children's blood pressure was taken along with height and weight. Analyses examined single exposure models with only Pb or As exposures over time as well as Pb and As in the same model. Adjustments in the final models were sex, age at time of assessment, child height and weight, and maternal smoking during pregnancy. Results are shown in the following table, adapted from Farzan et al. (2018, p. 21):

Exposure Measure	N	SBP β (95% CI)	SPB p-value	DBP β (95% CI)	DPB p-value
Prenatal Toenail Pb	257	0.58 (0.05, 1.11)	0.032	0.17 (-0.16, 0.49)	0.31
Postpartum Toenail Pb	285	-0.08 (-0.59, 0.44)	0.77	-0.14 (-0.44, 0.16)	0.36

The inconsistent results of this study show only one (among four possible) statistically significant and weak associations. By "weak" I mean that that "maternal prenatal toenail Pb…was associated with a 0.58 mmHg increase in children's systolic blood pressure (SBP). There were no associations observed between postpartum Pb and systolic (SBP) or diastolic blood pressure (DBP).

Cajan et al. (2018) is a follow-up study of the relationship between childhood lead exposure and hypertension in adults in a cohort of individuals exposed in the Baltic country of Kosovo. The investigators found no evidence of an association. See Cajan et al. (2018, p. 6):

> "…earlier measures of BPb were not significantly associated with either BP measure.."

By "earlier measures of BPb" the authors are referring to childhood lead exposures. The results of this study refute any claim that the Flint, Michigan plaintiffs are at an increased risk of hypertension in adults.

Sanders et al. (2018) is a study of the relationship between gestational age and blood pressure in children and the extent to which lead exposure could mediate that relationship. The authors identified mother-child pairs participating in a Mexican longitudinal cohort study. Women were recruited in their second trimester (2007-2011) and followed through birth to when the children were 4-6 years of age. Demographic and medical information was collected including maternal age, socioeconomic status, gestational age at delivery, birth weight, age at BP measurement, height and sex. Women were excluded from the study if they were daily alcohol users. Maternal smoking was determined as well as household environmental tobacco smoke. Lead exposure was measured from maternal venous samples at the second trimester visit.

The authors examined the relationship between continuous gestational age and BP at 4-6 years using linear and non-linear models adjusting for child sex, height, age at time of BP measurement, maternal socioeconomic status, and environmental tobacco smoke exposure inside the home. The calculated beta coefficients represent the change in BP (mm Hg) per week change in gestation. In addition, the authors identified a cutpoint of 2.5 µg/dL of lead in order to examine whether the relationship between gestational age and BP differed below versus above that cutpoint.

Results revealed that for those mother-child pairs where the maternal blood level at 2nd trimester was ≥ 2.5 µg/dL, the systolic blood pressure (SBP) was 1.6 (95% CI: 0.3-2.9) mm Hg higher per each week reduction in gestational age among children born before 37 weeks. Among the children born after 37 weeks, the SBP was 0.9 (95% CI: 0.2-1.6) mm Hg higher.

The authors conclude (Sanders et al. 2018, p. 470):

> "We found that higher maternal BLLs in pregnancy modified the association between lower gestational age and increased BP in children, whereby children with combined prematurity and maternal BLL in pregnancy ≥ 2.5 µg/dL had higher BP at age 4-6, than children who did not have both characteristics."

Zachariah et al. (2018) is the report of a study of cross-sectional data examining the relationship between lead exposure and trends in blood pressure and obesity. The authors note that obesity is a known risk factor for increased blood pressure in children (and adults). They also note that obesity prevalence has increased in the U.S. The study was designed to examine the extent to which blood pressure trends in children—which should rise as a result of increased obesity—were modified by blood lead levels. The data for the study came from the National Health and Nutrition Examination Surveys (NHANES), 1976-2008. The authors identified over 13,000 children aged 8-17 with anthropometric data,

BP, and laboratory data. Multivariable adjusted survey regression was used to examine how BP measures changed (or not) over time relative to increasing obesity and lead levels. The authors write:

> "As obesity prevalence rose from 5.3% to 24.5%, age-sex adjusted SBP was flat (-0.01[95% CI:-0.06,0.4]mmHg/yr, p=0.8), while DBP declined (-0.28 [95% CI: -0.32,-0.24]mmHg/yr, p<0.001). Accounting for blood lead concentration attenuated the DBP decline."

Importantly, this cross-sectional data does not permit the inference that blood lead was causally related to blood pressure.

Yao et al. (2020) is the report of a study of cross-sectional data examining the relationship between lead exposure and blood pressure using the NHANES datasets (2007-2016). Outcome variables were systolic BP (SBP), diastolic blood pressure (DBP), and high BP status (defined as using hypertension medication or a diagnosis of hypertension). Study participants were children and adolescents 8-17 years (n = 7076). The authors found no association between lead levels and either SBP or DBP. See Yao et al. (2020, p. 3):

> "In the multivariable-adjusted linear regression analysis, no significant association between either systolic BP or diastolic BP and blood lead levels was found when all participants were entered in the model."

Summary of the Blood Lead and Blood Pressure Studies (2018-present)

Studies of blood lead levels and blood pressure in children and adolescents during the years 2018-2020 reveal inconsistent and largely negative results. These studies do not change the overall conclusion that it has not been established that elevated blood lead levels causally influence blood pressure measurements in children and adolescents. Furthermore, there is no evidence that blood lead levels in children causes hypertension much less the health consequences of hypertension.

An effort to apply the Bradford-Hill criteria to this evidence will fail. Some studies are cross-sectional and cannot therefore contribute to a causal analysis. Furthermore, the results across all studies are inconsistent, the measured "effects" are weak, and there is no clear evidence of a dose response relationship. Indeed, it is not clear that an association exists "beyond that which we would attribute to chance" (quoting Hill, 1965) between lead and increased blood pressure in children. In addition, there is no observed association between lead and hypertension or the "health consequences of hypertension" in children. In the absence of an association, applying the Bradford-Hill criteria is scientifically inappropriate.

References

Agency for Toxicological Substances and Disease Registry (ATSDR). Toxicological Profile for Lead. August, 2020. Department of Health and Human Services.

Ahn J, Kim NS, Lee BK, et al. Association of blood pressure with blood lead and cadmium levels in Korean adolescents: Analysis of data from the 2010-2016 Korean National Health and Nutrition Examination Survey. J Korean Med Sci 2018;33(44):e278.

Camaj PR, Graziano JH, Preteri E, et al. Long-term effects of environmental lead exposure on blood pressure and plasma soluble cell adhesion molecules in young adults: A follow-up study of a prospective cohort in Kosovo. J Environ Pub Health 2018;3180487.

Dong Y, Song Y, Zou Z, et al. Updates to pediatric hypertension guidelines: influence on classification of high blood pressure in children and adolescents. J Hypertension 2019;37:297-306.

EPA. United States Environmental Protection Agency. Integrated Scientific Assessment for Lead. EPA/600/R-10/075F/June 2013. Office of Research and Development, National Center for Environmental Assessment, Research Triangle Park, North Carolina.

Farzan SH, Howe CG, Chen Y, et al. Prenatal lead exposure and elevated blood pressure in children. Environ Int 2018;121(Pt. 2):1289-96.

Flynn JT, Kaelber DC, Baker-Smith CM, et al. Clinical Practice Guideline for Screening and Management of High Blood Pressure in Children and Adolescents. Pediatrics 2017;140(3):e20171904.

Flynn JT, Falkner BE. New clinical practice guideline for the management of high blood pressure in children and adolescents. Hypertension 2017;70:683-6.

Gerr F, Letz R, Stokes L, et al. Association between bone lead concentration and blood pressure among young adults. Am J Indust Med 2002;42:98-106.

Gump BB, Stewart P, Reihman J, et al. Prenatal and early childhood blood lead levels and cardiovascular functioning in 9 ½ year old children. Neurotoxicol Teratol 2005;27:655-65.

Gump BB, MacKenzie JA, Bendinskas K, et al. Low-level Pb and cardiovascular responses in acute stress in children: The role of cardiac autonomic regulation. Neurotoxicol Teratol 2011;33:212-9.

Sanders AP, Svensson K, Gennings C, et al. Prenatal lead exposure modifies the effect of shorter gestation on increased blood pressure in children. Environ Int 2018;120:464-71.

Siddiqui S, Malatesta-Muncher R. Hypertension in children and adolescents: A review of recent guidelines. Pedatr Ann 2020;49:e250-7.

Yao B, Lu X, Xu L, et al. Relationship between low-level lead, cadmium and mercury exposures and blood pressure in children and adolescents aged 8-17 years: An exposure-response analysis of NHANES 2007-2016. Sci Total Environ 2020;726:138446.

Zachariah JP, Wang Y, Penny DJ, et al. Relation between lead exposure and trends in blood pressure in children. Am J Cardiol 2018;122:1890-5.

Zhang A, Hu H, Sanchez BN, et al. Association between prenatal lead exposure and blood pressure in children. Environ Health Perspect 2012;120:445-50.

Appendix H

An Assessment of the Proposed Relationship Between Lead Exposure in Children and Cardiovascular Disease in Adults who were Exposed as Children

An initial review of governmental reports (i.e., ATSDR 2020; NTP 2012; EPA 2013) was conducted to identify primary sources. Once these documents and the primary sources were reviewed, literature searches were conducted on several health outcomes using the library databases of PubMed and Web of Science. Specifically, this review was conducted to identify longitudinal studies that evaluated childhood blood lead levels and development of cardiovascular diseases in adulthood. Details regarding the searches and the health outcomes are as follows:

PubMed search terms: ("Cardiovascular diseases"[Mesh]) AND ("Lead/Blood"[Mesh]) Results: n = 221
Web of Science search terms: ("blood lead") AND TS= ("cardiovascular disease" OR "CVD" OR "stroke")
Results: n = 211

After removal of duplicates, there were 389 publications examined. No longitudinal studies as defined above were identified.

A claim of general causation that exposure to lead in childhood leads to (whether "associated with" or "causes") cardiovascular disease in adulthood is absent any epidemiologic evidence.

Appendix I

An Assessment of the Proposed Relationship Between Lead Exposure in Children and Essential Tremor in Adults who were Exposed as Children

An initial review of governmental reports (i.e., ATSDR 2020; NTP 2012; EPA 2013) was conducted to identify primary sources. Once these documents and the primary sources were reviewed, literature searches were conducted on several health outcomes using the library database of PubMed. Specifically, this review was conducted to identify longitudinal studies that evaluated childhood blood lead levels and development of essential tremor in adulthood. Details regarding the searches and the health outcomes are as follows:

PubMed search terms: "Essential Tremor" and "Lead Exposure" and "Epidemiology" (n = 6)
No longitudinal studies as defined above were identified.

A claim of general causation that exposure to lead in childhood leads to (whether "associated with" or "causes") essential tremor in adulthood is absent any epidemiologic evidence.